Structural and Functional Analysis of Virulence-Associated Listerial Surface Proteins

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades einer

Doktorin der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

Dissertation

von Maike Bublitz
aus Wolfenbüttel

1. Referent:	Honorarprofessor Dr. Dirk Heinz
2. Referent:	Professor Dr. Michael Steinert
eingereicht am:	05.11.2007
mündliche Prüfung (Disputation) am:	20.12.2007

Druckjahr 2008

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Tagungsbeiträge

Bublitz, M., Heinz, D. W. & Schubert, W. D.: Structural Basis for Autoinhibition of an Autolysin involved in *Listeria monocytogenes* Pathogenicity. (Poster), Murnau Conference on Structural Biology of Disease Mechanisms, Murnau (2007).

Bublitz, M., Holland, C., Heinz, D. W. & Schubert, W. D.: Crystal Structure of InlJ': A Leucine-Rich Repeat Protein with a Cysteine Ladder. (Poster), Jahrestagung der Deutschen Gesellschaft für Kristallographie, Bremen (2007). Poster-Preis.

Bublitz, M., Heinz, D. W. & Schubert, W. D.: Peptidoglycan and Pathogenesis – Characterization of a Listerial Autolysin. (Vortrag) 9th Heart of Europe Bio-Crystallography Meeting, Teistungenburg (2006).

Bublitz, M., Heinz, D. W. & Schubert, W. D.: Structure and Function of an Autolysin involved in the Pathogenicity of *Listeria monocytogenes.* (Poster) 8th International School on the Crystallography of Biological Macromolecules, Como, Italien und Jahrestagung der Deutschen Gesellschaft für Kristallographie, Freiburg (2006).

Bublitz, M., Stradal, T., Heinz, D. W. & Schubert, W. D.: The Internalin Family of Listerial Surface Proteins: Analysis of LRR-Proteins involved in Infection. (Poster) Murnau Conference on Structural Biology of Molecular Recognition, Murnau (2005).

Für Benno Laumert,
der viel zu früh das getan hat, was nur die Besten tun.

Contents

Abbreviations

aa	Amino acid(s)
Å	Ångstrøm (1 Å = 0.1 nm)
A_{λ}	Absorption at the wavelength λ in nm (equivalent to $OD_{\lambda)}$
ActA	Gene product of *actA* (lmo0204) from *L. monocytogenes* EGD-e
AEBSF	4-(-Aminoethyl) benzenesulphonyl fluoride hydrochloride
APS	Ammonium persulphate
aut	Gene lmo1076 from *L. monocytogenes* EGD-e, encoding the autolysin Auto
Δaut	*aut* deletion mutant of *L. monocytogenes* EGD-e
Auto	Gene product of *aut* (lmo1076) from *L. monocytogenes* EGD-e
Auto$_{act}$	Amino acids 84-243 of Auto, comprising the active catalytic domain
bCC-LRR	Bacterial cysteine-containing leucine-rich repeat
BESSY	Berliner Elektronenspeicherring-Gesellschaft für Synchrotron-strahlung
BHI	Brain Heart Infusion (medium)
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
C-	Carboxy-
c-Slt70	Catalytic domain of the soluble lytic transglycosylase Slt70 from *E. coli*
CAPS	*N*-Cyclohexyl-3-aminopropanesulphonic acid
CID	Collision-induced dissociation
D-*iso*Glu	D-*iso*-glutamate
DESY	Deutsches Elektronen-Synchrotron
DLS	Dynamic light scattering
DMEM	Dulbecco's modified Eagle's medium
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleosid-triphosphate
DTNB	5,5'-Dithio-bis(2-nitrobenzoic acid)
DTT	Dithiothreitol
E. coli	*Escherichia coli*
E-64	L-*trans*-epoxysuccinyl-leucyl amido(4-guanidio)butane
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EMBL	The European Molecular Biology Laboratory
ESI	Electrospray ionisation
EtOH	Ethanol
FCS	Foetal calf serum
GEWL	Goose egg-white lysozyme
GFP	Green fluorescent protein
GH19, 23 etc.	Glycoside hydrolase family (according to the CAZY database)
GPC	Gel permeation chromatography
GST	Glutathione S Transferase

HCOOH	Formic acid
HEWL	Hen egg-white lysozyme
HF	Hydrofluoric acid
His_6	6 consecutive histidine residues
HRP	Horseradish peroxidase
Hyg	Threo-3-hydroxy-glutamic acid
IEC	Ion exchange chromatography
InlA, -F, -G	Internalin family members (Internalin, Internalin F, Internalin G) from *L. monocytogenes* EGD-e
InlA'	Internalin domain of InlA (lmo0433) from *L. monocytogenes* EGD-e, comprising residues 36-495
InlG'	Internalin domain of InlG (lmo0262) from *L. monocytogenes* EGD-e, comprising residues 27-277
InlJ'	Internalin domain of InlJ (lmo2821) from *L. monocytogenes* EGD-e, comprising residues 34-508
InlX	Any of the internalin constructs generated in this work
IPTG	Isopropyl β-D-thiogalactopyranoside
K_M	Michaelis-Menten constant
K_S	Substrate inhibition constant
LB	Luria Bertani
LIPI-1	*Listeria* pathogenicity island 1
LL-DAP	L,L-diaminopimelic acid
LLO	Listeriolysin O
LRR	Leucine rich repeat
LTA	Lipoteichoic acid
MAD	Multiple anomalous dispersion
MALDI	Matrix assisted laser desorption / ionisation
*m*DAP	*Meso*-diaminopimelic acid
MeOH	Methanol
Mpl	Metalloprotease, gene product of *mpl* (lmo0203) from *L. monocytogenes* EGD-e
MS	Mass spectrometry
MW	Molecular weight
N-	Amino terminal
n.d.	Not deposited
NAG	*N*-acetylglucosamine
NAM	*N*-acetylmuramic acid
NCS	Non-crystallographic symmetry
NEAA	Non-essential amino acids
no.	Number
o.n.	over night
OD_λ	Optical density at the wavelength λ in nm (equivalent to A_λ)
PAGE	Polyacrylamide gelelectrophoresis
PBS	Phosphate buffered saline
PBST	PBS containing 0.1 % (v/v) TWEEN-20
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PG	Peptidoglycan

PlcA	Phosphatidylinositol-specific phospholipase c (PI-PLC), gene product of *plcA* (lmo0201) of *L. monocytogenes* EGD-e
PlcB	Broad-range phospholipase c (PC-PLC), gene product of *plcB* (lmo0205) of *L. monocytogenes* EGD-e
PrfA	Positive regulatory factor A, gene product of *prfA* (lmo0200) from *L. monocytogenes* EGD-e
pro-Auto	Amino acids 27-243 of Auto, comprising the inhibitory N-terminal plug (aa 27-83) and the catalytic domain (aa 84-243)
psi	Pounds per square inch
PVDF	Polyvinylidene difluoride
r.m.s.d.	Root mean square deviation
RNA	Ribonucleic acid
rpm	Rotations per minute
SDS	Sodium dodecyl sulphate
SeMet	Selenomethionine
T_m	Melting temperature
TEMED	N, N, N', N'-Tetramethylethylenediamine
TEV-protease	Catalytic domain of the 'nuclear inclusion a' protein encoded by the tobacco etch virus
TOF	Time of flight
TNB	2-nitro-5-mercaptobenzoic acid
Tris	Tris-(hydroxymethyl)aminomethane
wt	Wild type

Summary

The cell wall of the Gram-positive human pathogen *Listeria monocytogenes* is decorated with numerous surface-associated proteins, a subset of which constitute virulence factors crucial for host-pathogen interactions such as adhesion, invasion and responding to the host immune system. The cell wall is continuously remodelled allowing its composition to be optimized *inter alia* for successful infection. This process, however, requires the tightly regulated activity of autolytic and cell wall-synthesizing enzymes. The autolysin Auto has recently been identified as a novel virulence factor in *L. monocytogenes*, indispensable for host cell invasion. In this work, the lytic activity and substrate specificity of Auto has been characterized, determining it to be an *N*-acetylglucosaminidase capable of hydrolyzing a wide range of peptidoglycan types. Initially autoinhibited pro-Auto is activated by proteolytic cleavage, removing 57 N-terminal amino acids and freeing the active site. The crystal structure of the autoinhibited catalytic domain of Auto (pro-Auto), solved in this study, provides the first structural data for the large glycoside hydrolase (GH) family 73. It identifies Auto as belonging to a novel lysozyme class, with a fold related to that of goose-type lysozymes and soluble lytic transglycosylases. However, Auto is unique to this superfamily of enzymes in possessing an N-terminal autoinhibitory α-helix that blocks the substrate-binding cleft. This feature allows its spatiotemporal activation during bacterial infection, providing the basis for the dynamic adaptation of its virulence-promoting surface proteome, presumably during the phagosomal step of infection.

Furthermore, the crystal structures of the internalin domains of InlJ and InlG have been determined at 2.5 and 2.05 Å, respectively. Both are members of the virulence-associated internalin family of listerial cell-surface proteins, characterized by a common N-terminal domain. Whereas InlG is the smallest cell-wall bound internalin with only 5 leucine-rich repeats (LRRs), InlJ bears 15 LRRs, the same number as InlA, the prototypical internalin family member. The LRRs of InlJ differ from those of other internalins in that each repeat consists of 21 residues instead of 22, and by replacing one LRR-defining hydrophobic residue by a conserved cysteine. These cysteines stack to form an intramolecular ladder and regular hydrophobic interactions in consecutive repeats. Analysis of the curvature, twist and lateral bending angles of InlJ and InlG, and comparison of these to several other leucine-rich repeat proteins, provides a systematic, geometric analysis of LRR-protein structures. This has been implemented in an online-tool (http://bragi2.helmholtz-hzi.de/Angulator/), making it available to the scientific community. The analysis indicates that both cysteine and asparagine ladders serve to stabilize LRR proteins and supports the assumption, that internalins have evolved as virulence-associated protein-protein interaction tools, with the key residues being located on their concave surfaces.

1 Introduction

Despite the discovery of antibiotics early in the 20th century (Fleming, 1929), and continuous efforts to develop novel therapies and vaccines, infectious diseases are still responsible for a large proportion of deaths worldwide (Fig. 1-1). A lack of sanitary facilities and financial resources in less-developed countries account for the majority of these deaths. Industrialized societies, however, also need to address the rising threat of antibiotic resistances in bacterial pathogens.

According to the World Health Organization (WHO), food-borne diseases continue to be a widespread and growing public health problem, both in developed and developing countries.

The economic burden created by food contamination is enormous, with estimated medical costs and lost productivity of 35 billion US$ (1997) in the USA alone (WHO, Fact sheet no. 237).

Intense research on the molecular mechanisms underlying pivotal host-pathogen interactions is needed to provide the framework for the development of novel antimicrobial drugs. The success of new strategies for a global combating of microbial infections in the future will strongly depend on the availability of high-resolution structures of virulence factors and host proteins involved in disease.

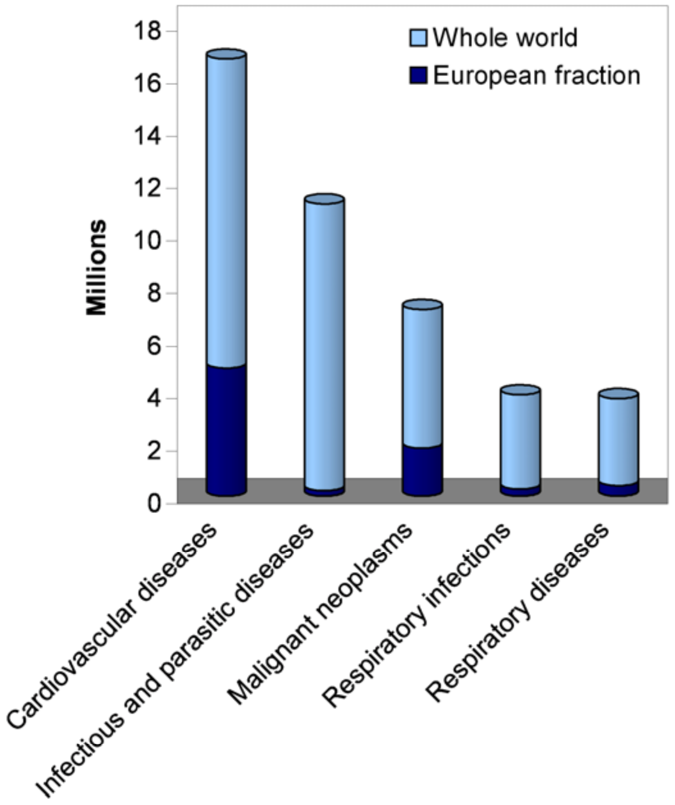


Fig. 1-1: Global causes of death (Global Burden of Disease Estimates 2002, WHO)

1.1 *Listeria monocytogenes* and listeriosis

Listeria monocytogenes is a ubiquitous soil and water bacterium. Together with *Lactobacilli*, *Listeriae* phylogenetically belong to the subdivision of Gram-positive, non-sporulating, low GC content eubacteria. *Listeriae* are facultatively anaerobic and particularly tolerant towards low pH, low temperatures, and high salt concentrations (Seeliger & Jones, 1986). Two of the currently known six *Listeria* species are potentially pathogenic: *L. monocytogenes* and *L. ivanovii* (Glaser *et al*., 2001). Its wide distribution and ability to grow under harsh environmental conditions make *L. monocytogenes* an important food-borne human pathogen (Altekruse *et al*., 1997). Ready-to-eat products such as soft cheese, salami and unpasteurized milk products are possible sources for *L. monocytogenes*, since refrigeration does not sufficiently inhibit bacterial growth (McLauchlin *et al*., 2004).

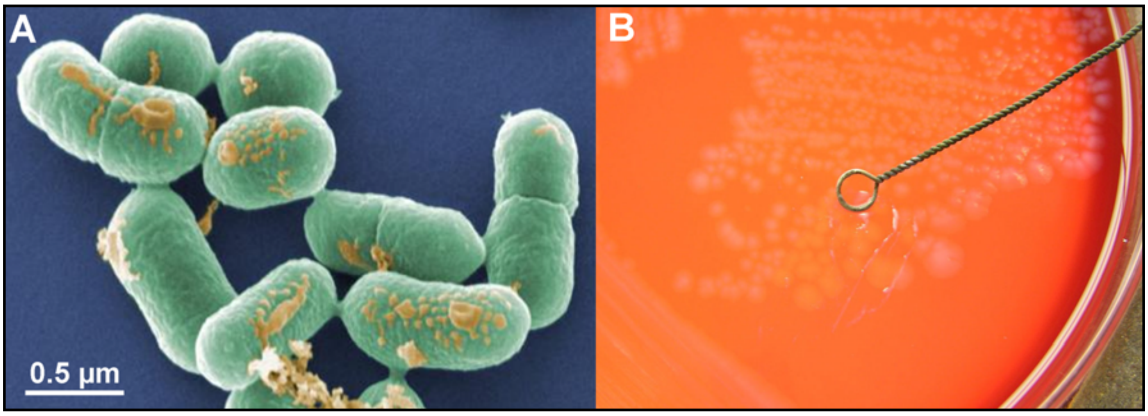


Fig. 1-2: *L. monocytogenes*. (A) Rod-shaped bacterial cells magnified 50 000-fold (source: Okapia KG). **(B)** Colonies of *L. monocytogenes* growing on a blood-agar plate (source: R. Buxton, www.microbelibrary.org).

Despite frequent human exposure to these bacteria, severe listerial infections are rare, because these generally only occur in immunosuppressed individuals, such as neonates, the elderly, pregnant women or AIDS patients. In healthy individuals, the immune system rapidly clears the infection, leading to transient symptoms such as gastroenteritis (Hof, 2001). When a systemic infection does occur, the mortality rate of 20-30 %, primarily due to listerial meningitis, by far exceeds that of other bacterial infections (Gellin & Broome, 1989; Mead *et al*., 1999). According to the Robert-Koch Institute, Berlin, the number of reported cases of listeriosis has steadily increased over the last few years. The true number of outbreaks is, however, probably underestimated, due to frequent incorrect diagnoses. In 2005, 519 cases of listeriosis were reported in Germany, roughly two thirds (346) of these in patients older than 60 years (Koch & Stark, 2006).

Listeriosis is mostly contracted by consuming contaminated food. Some bacteria that reach the gastrointestinal tract are unspecifically taken up by M-cells that cover the intestinal Peyer's patches and actively invade intestinal epithelial cells at the cell extrusion zone of villus tips (Pentecost *et al.*, 2006; Wollert *et al.*, 2007b). Once inside a phagosome, bacteria escape into the cytoplasm of the host cells (Gaillard *et al.*, 1987). Here they proliferate, move by reorganizing the host actin cytoskeleton, disseminate directly to neighbouring cells and eventually reach the lymph and blood systems, allowing them to spread throughout the host organism. Manifestations of listeriosis include sepsis, invasion of liver and spleen, meningitis, encephalitis, miscarriage and eventually death (Vazquez-Boland *et al.*, 2001b). These severe stages of the infection, however, only occur if the host's cellular immune response is attenuated, and if the pathogen is not identified and treated in time.

Individual steps of listerial infections have been studied in remarkable molecular detail, allowing each step of infection to be assigned to individual virulence factors (Fig. 1-3).

Key players during infection are the invasion proteins InlA and InlB that mediate uptake into a range of host cell types, depending on the presence of the respective receptors E-cadherin and c-Met (Mengaud *et al.*, 1996; Shen *et al.*, 2000).

Escape from the primary phagosome is mainly accomplished by the secreted pore forming toxin listeriolysin (LLO), aided by the phosphatidylinositol-specific phospholipase PlcA (Camilli *et al.*, 1993; Gaillard *et al.*, 1987). ActA, an activator of the Arp2/3 complex, induces actin polymerization at one pole of the rod-shaped bacterium, leading to randomized movement through the cytosol (Domann *et al.*, 1992; Kocks *et al.*, 1992). The bacteria proliferate and spread to neighbouring host cells, lysing the double membrane of the secondary phagosome with LLO and PlcB, which is activated upon phagosome acidification by the metalloprotease Mpl (Marquis *et al.*, 1995; Smith *et al.*, 1995). The virulence factors LLO, PlcA, PlcB, Mpl and ActA, as well as their master expression regulator PrfA, are all encoded in a central virulence gene cluster, LIPI-1 (Vazquez-Boland *et al.*, 2001a).

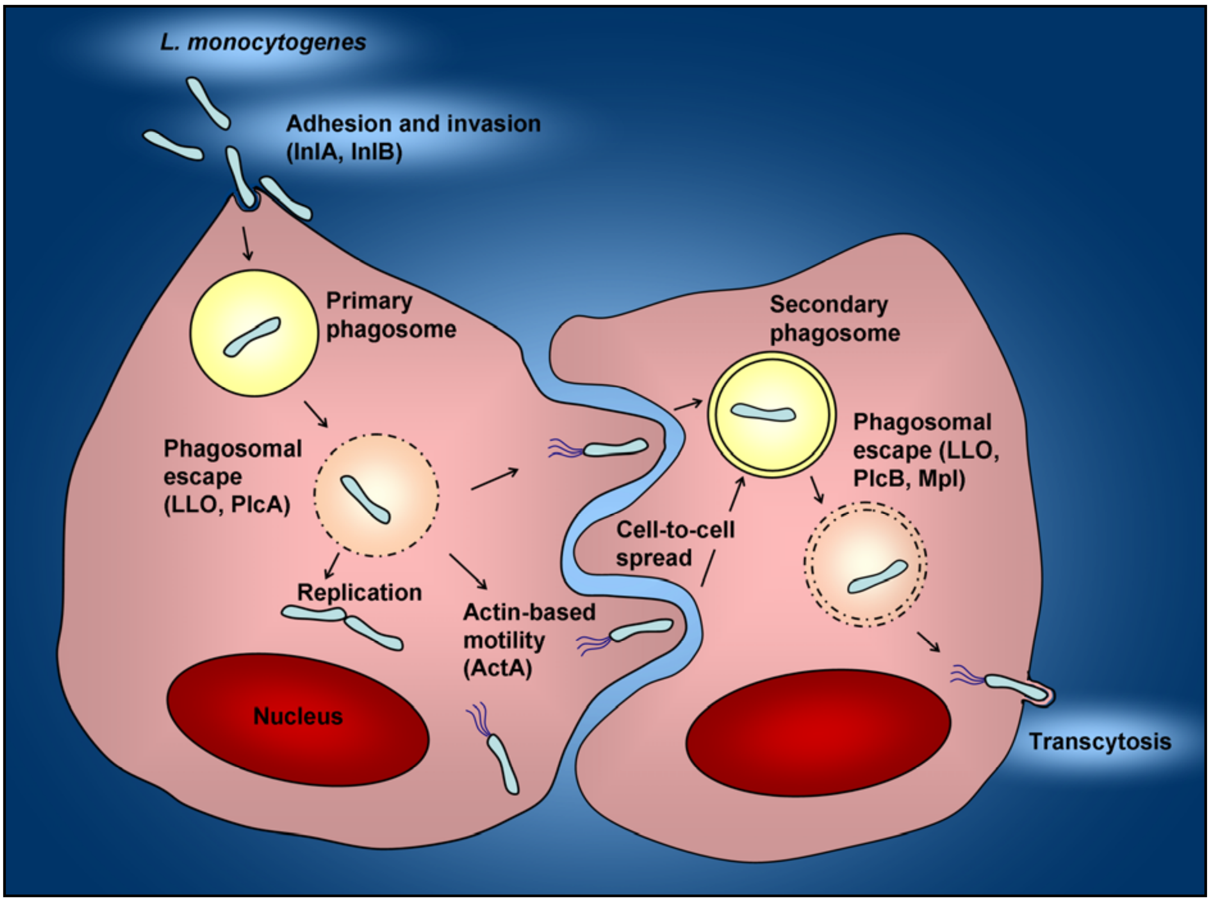


Fig. 1-3: Infection cycle and major virulence factors of *Listeria monocytogenes* (adapted from Pamer, 2004). Active host cell invasion is mediated by InlA and InlB, each of which specifically interacts with its particular receptor - the epithelial cell-cell-adhesion molecule E-cadherin and the hepatocyte growth factor receptor c-Met. Acidification of the resulting phagosome triggers activation and release of the pore-forming toxin LLO and the phospholipase PlcA. Polar accumulation of ActA at the bacterial surface promotes host-actin polymerization, allowing bacteria to move through the cytosol and to infect neighbouring cells. A secondary phagosome is formed, the double membrane of which is again perforated by LLO and Mpl-activated PlcB.

1.2 Surface-located and secreted virulence factors mediate intracellular infection

The host cell cytoplasm is a beneficial habitat for a bacterial pathogen, offering temporary protection from the humoral immune response. *L. monocytogenes* has adapted to the sophisticated tasks of host cell invasion, phagosome escape, intracellular motility and intercellular transmission by the evolution of an effective set of virulence genes.

A comparison of the genomes of *L. monocytogenes* and apathogenic *L. innocua* reveals that the two species differ mainly in their surface proteome. 4.7 % of all genes in *L. monocytogenes* encode surface proteins, and of these more than 22 % have no orthologue in *L. innocua* (Glaser *et al*., 2001).

Direct interactions between pathogen and host cell surfaces can only occur through individual molecules directly bound to the respective cell surface or through secreted factors. Accordingly, the more complex and versatile the manipulations of host functions by the pathogen are, the larger is the benefit to the pathogen to use of a multitude of specialized surface proteins. The dynamic reorganization of the bacterial cell surface hence appears essential to ensure the spatiotemporal orchestration of surface proteins during individual steps of an infection.

In *L. monocytogenes,* four mechanisms have been identified to anchor proteins to the cell wall:

i) sortase-dependent covalent linkage to the cell wall peptidoglycan through an LPXTG motif,
ii) non-covalent attachment to lipoteichoic acid mediated by GW domains,
iii) membrane anchoring of hydrophobic tail proteins,
iv) insertion into the membrane lipid bilayer by lipoproteins (Fig. 1-4, Cabanes *et al*., 2002).

The dynamic modification of the protein composition at the cell surface requires an efficient cell wall turnover machinery. Studying virulence-associated surface proteins as well as components of the dynamic cell wall metabolism serves to reveal the molecular processes shaping the interactions between bacterial and host cells.

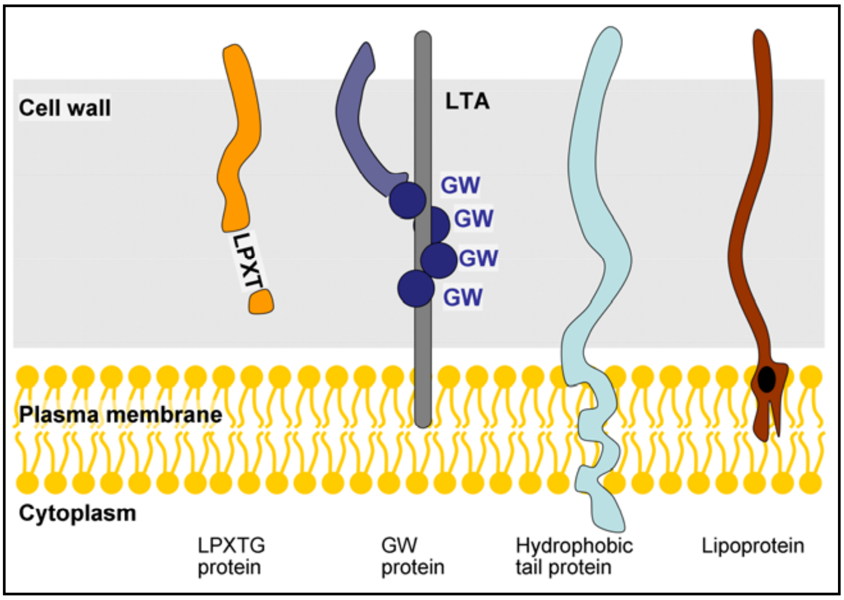


Fig. 1-4: The four major types of surface proteins found in *L. monocytogenes*: proteins with a C-terminal LPXTG-motif, proteins containing GW-modules, proteins anchored by a hydrophobic tail and lipoproteins (adapted from Cabanes *et al*., 2002).

1.3 Autolysins

1.3.1 The Gram-positive cell wall

The bacterial cell wall murein constitutes a rigid envelope, essential in maintaining the cell shape, counteracting the cytoplasmic hydrostatic pressure (turgor) and supporting the flagellar apparatus. Murein is a peptidoglycan (PG) heteropolymer, with an aminosugar backbone of alternating β-(1,4)-linked *N*-acetylglucosamine (NAG) and *N*-acetylmuramic acid (NAM) residues, crosslinked by short peptides. These peptides predominantly contain L- and D-alanine, D-*iso*glutamate (D-glutamate linked to the following amino acid via the γ-carboxy group), and a diamino acid such as L-lysine or *meso*-diaminopimelic acid (*m*DAP). The diamino acid allows peptides to be cross-linked either directly or through additional residues. The aminosugar backbone is largely invariant among eubacteria, whereas the peptide compositions are species-specific and were used as taxonomic markers before ribotyping methods became available (Schleifer & Kandler, 1972).

The Gram-positive murein sacculus may be regarded as a single, large yet flexible three-dimensional macromolecule, highly crosslinked and several layers thick (Archibald *et al.*, 1993). Apart from its mechanical functions, PG is constantly remodelled, to adapt to the particular environment encountered by the bacterium. These cell wall dynamics are achieved through an endogenous set of PG synthesizing and hydrolyzing enzymes – the latter being known as autolysins (Shockman & Höltje, 1994).

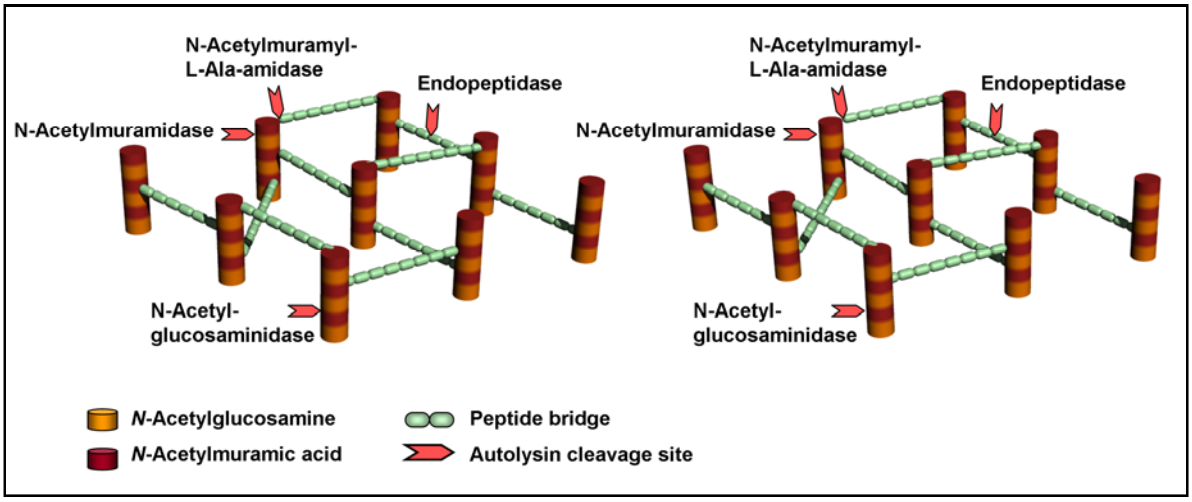


Fig. 1-5: Schematic stereo diagram of bacterial peptidoglycan. The three-dimensional network of peptide-cross-linked oligosaccharides constitutes a flexible but highly robust macromolecular sacculus. Cleavage sites of the four main autolysin classes are indicated by red arrowheads.

Alongside housekeeping functions such as cell growth, division, cell wall turnover, flagella assembly, competence for genetic transformation, sporulation and protein secretion, autolytic enzymes are also involved in antibiotic-induced autolysis, host cell adhesion and pathogenicity (Doyle *et al.*, 1988; Popowska, 2004; Smith *et al.*, 2000).

Autolysins are classified according to the bond in their PG substrate that they hydrolyze: *N*-acetylmuramidases and *N*-acetylglucosaminidases, respectively, cleave the NAM-(1,4)-NAG and the NAG-(1,4)-NAM glycosidic bond. Lytic transglycosylases cleave the same bond as *N*-acetylmuramidases, but generate a non-reducing 1,6-anhydro-*N*-acetylmuramic acid product. *N*-acetylmuramyl-L-alanine amidases (NAM-L-Ala amidases) cleave the amide bond between the lactyl moiety of NAM and L-alanine of the bridge peptide, whereas endopeptidases specifically cleave bonds within the peptide bridge. In *L. monocytogenes*, PG and the associated (lipo)teichoic acid polymers provide a docking site for surface proteins that interact with the host. Bacterial adhesion, host cell invasion and interaction with the host immune system all depend on a functional (i.e. temporally and spatially adapted) set of virulence factors located on the bacterial cell surface (Hamon *et al.*, 2006; Vazquez-Boland *et al.*, 2001b). Listerial PG is unusual amongst those of Gram-positive bacteria, by containing *m*DAP instead of the more widespread lysine, to mediate the cross-linking of peptide bridges (Schleifer & Kandler, 1972).

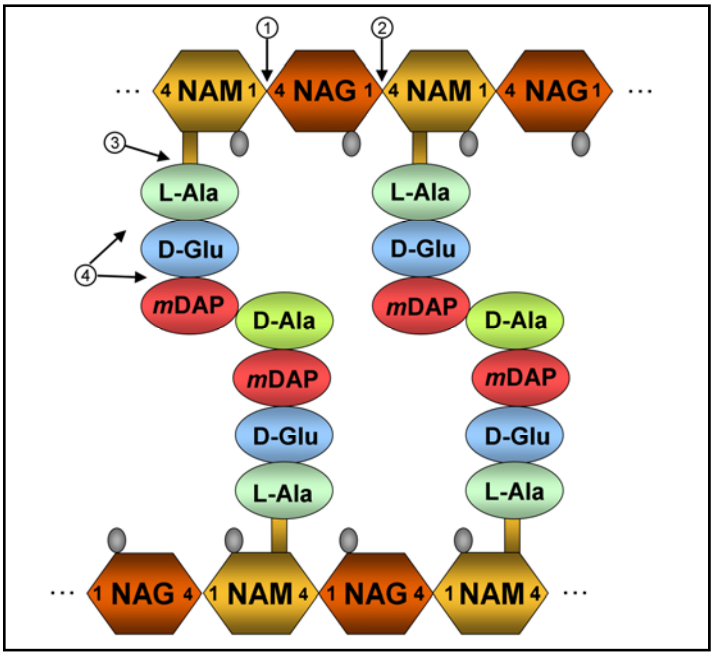


Fig. 1-6: Schematic representation of listerial peptidoglycan. The third residue of the peptide bridge, *meso*-diaminopimelic acid (*m*DAP), mediates direct cross-links between two peptides. Grey ellipses indicate acetyl moieties. Autolysin cleavage sites are indicated by numbered arrows: (1): *N*-acetylmuramidase or lytic transglycosylase, (2): *N*-acetylglucosaminidase, (3): *N*-acetylmuramyl-L-Ala amidase, (4): endopeptidase.

1.3.2 Autolysins in *L. monocytogenes*

L. monocytogenes encodes multiple autolysins and three additional phage-derived endolysins. P60, a cell wall associated autolysin encoded by the *iap* gene, is presumably involved in cell separation and invasion of certain host cells (Hess *et al.*, 1995; Kuhn & Goebel, 1989). P45 is a functionally uncharacterized 45-kDa homologue of p60, encoded by the gene *spl*. Ami is a 102 kDa autolysin with eight C-terminal GW-domains that ensure a tight non-covalent binding to cell wall lipoteichoic acids – similar to the four GW-domains of Auto (see 1.3.3) and in well-studied InlB (Braun *et al.*, 1997). This cell wall-anchoring domain promotes Ami adhesion to the cell wall, contributing to virulence (Milohanic *et al.*, 2001). Zymography identified NamA (or MurA) as a major autolytic cell-surface protein, its loss affecting autolysis and cell separation (Carroll *et al.*, 2003). FlaA constitutes the flagellum in motile *Listeriae* and has autolytic activity (Popowska & Markiewicz, 2004). The phage-derived endolysins Ply118, Ply500 and Ply511 act specifically on listerial PG and can induce rapid cell lysis when applied exogenously (Loessner *et al.*, 1995).

However, none of the endogenous autolysins of *L. monocytogenes* has as yet been experimentally classified with respect to their PG cleavage sites.

1.3.3 The autolysin Auto

A comparison of the genomes of *L. monocytogenes* and the non-pathogenic *L. innocua*, recently identified the novel autolysin Auto (lmo1076) as the only autolysin absent from the latter (Cabanes *et al.*, 2004). Auto was shown to be required for entry of *L. monocytogenes* into numerous cell lines, including Vero, Hep-2, Caco-2, L2, and GPC16 cells. Since adhesion was not observed to be dependent on Auto, it probably contributes to active host cell invasion, comprising InlA- and InlB-mediated phagocytosis and escape from the phagosome into the cytoplasm. Consistent with the invasion assays using cultured cell lines, an *aut* deletion mutant strain of *L. monocytogenes* had a reduced invasiveness *in vivo*, as shown in orally infected guinea pigs and intravenously infected BALB/c mice (Cabanes *et al.*, 2004).

Like the other known listerial autolysins, Auto is a surface-located enzyme with a modular architecture. Its 572 amino acid sequence comprises an N-terminal export signal (residues 1-26), a PG hydrolase domain (residues 89-243), and four C-terminal GW modules (residues 244-572). The C-terminal region is homologous to the non-covalent cell wall anchoring motifs of the invasion protein InlB and the autolysin Ami (Braun *et al.*, 1997; Marino *et al.*, 2002; Milohanic *et al.*, 2001). The function of residues 27 to 88, constituting the native N-terminus after cleavage of the secretion signal upon translocation of Auto, was unknown at the beginning of this work.

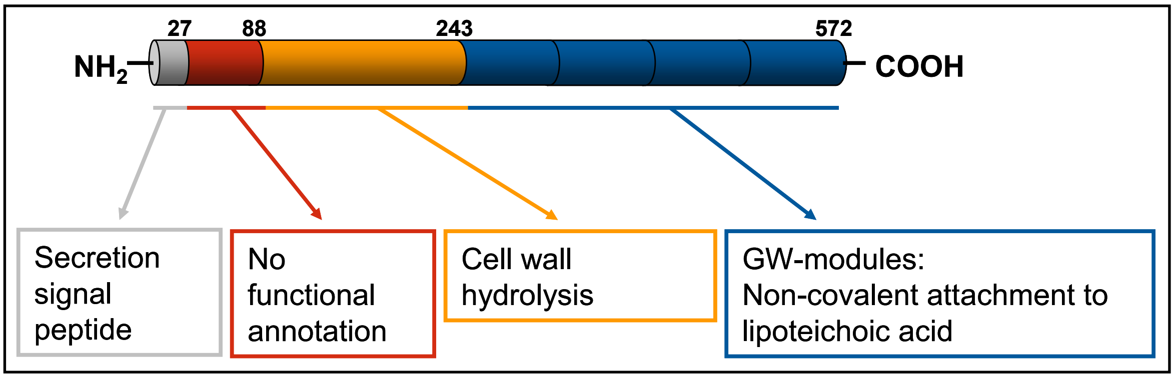


Fig. 1-7: Domain organization of Auto. Auto has a modular architecture, the cell wall hydrolysis domain being N-terminally flanked by a region of hitherto unassigned function, and C-terminally by four cell wall anchoring GW modules.

1.4 Internalins: Surface-located leucine-rich repeat proteins

Internalin-related proteins constitute a multigene family exclusive to *Listeriae*, defined by a tripartite N-terminal internalin domain. An α-helical cap at the N-terminus is followed by variable numbers of leucine-rich repeats (LRRs), the last of which passes into the β-sandwich of an Ig-like fold (Fig. 1-8, Schubert *et al.*, 2001). *L. monocytogenes* internalins are surface-located proteins, the only exception being the secreted InlC (Vazquez-Boland *et al.*, 2001b). The internalins of *L. ivanovii*, by contrast, are all secreted (Engelbrecht *et al.*, 1998). Most internalins of *L. monocytogenes* are covalently anchored to the bacterial cell wall via a C-terminal LPXTG-motif. Structurally, the LRR-region forms a solenoid-like superhelix, the overall bending and twisting geometry of which is dictated by the consensus sequence of the particular protein and by the amino acid sequence of each particular repeat (Kajava, 1998). The availability of extended surfaces combined with a rigid scaffold make LRR-proteins ideal for protein-protein interactions, a property in common with other repeat proteins (Djinovic-Carugo *et al.*, 2002; Goebl & Yanagida, 1991; Kobe & Deisenhofer, 1994; Mosavi *et al.*, 2004). LRR-proteins are found in plants, yeast, numerous metazoans and in bacteria (Buchanan & Gay, 1996). Amongst bacteria, a family of 19 internalin-family genes is unique to the listerial genome (Cabanes *et al.*, 2002; Vazquez-Boland *et al.*, 2001b).

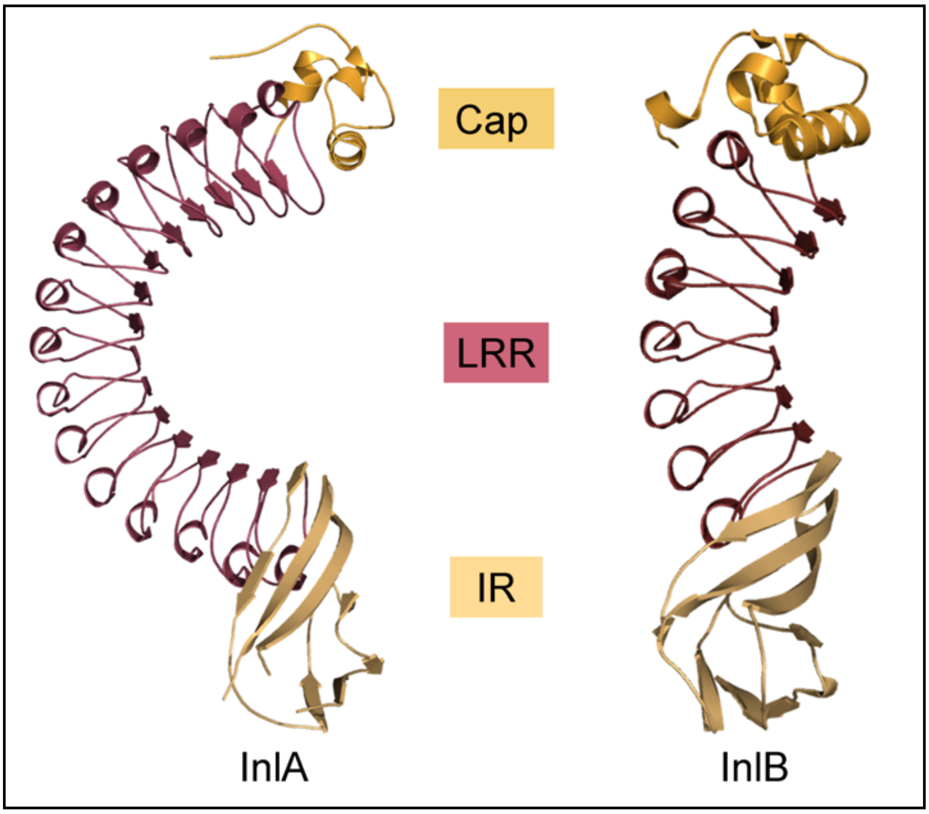


Fig. 1-8: Typical internalin domain architecture exemplified by InlA and InlB. The number of leucine-rich repeats dictates the overall curvature of the LRR region.

Two listerial internalins, InlA and InlB, have been particularly well studied both functionally (Dramsi *et al.*, 1995; Gaillard *et al.*, 1991) and structurally (Fig. 1-8, Marino *et al.*, 1999; Schubert *et al.*, 2002). Both are major virulence factors of *L. monocytogenes* and in particular, both are invasion proteins that enable bacterial uptake by normally non-phagocytic host cells (Dramsi *et al.*, 1995; Gaillard *et al.*, 1991; Hamon *et al.*, 2006). The crystal structures of the functional domains of InlA (Schubert *et al.*, 2002; Wollert *et al.*, 2007a), InlB (Marino *et al.*, 1999; Schubert *et al.*, 2001), InlC (Ooi *et al.*, 2006), InlE (Schubert & Heinz, 2003) and InlH (Schubert *et al.*, 2001) have been published. Crystal structures of InlA in complex with its human receptor E-cadherin (Schubert *et al.*, 2002), and of InlB in complex with its receptor c-Met (Niemann *et al.*, 2007) have also been reported.

Apart from InlA and InlB, other internalins, such as InlC, -G, -H, -E and -J have been found to be involved in infection to a greater or lesser extent (Ooi *et al.*, 2006; Raffelsbauer *et al.*, 1998; Sabet *et al.*, 2005), however, their explicit functions as well as their putative host receptors remain unknown.

1.4.1 InlJ

InlJ, encoded by the gene *inlJ* (lmo2821) is a recently identified member of the internalin family, sharing its defining features: the N-terminal internalin domain (consisting of an α-helical cap, a LRR region and an Ig-like interrepeat domain) and a C-terminal LPXTG-motif (Sabet *et al*., 2005). The LPXTG-motif is followed by a stretch of hydrophobic residues (transmembrane α-helix) and four basic amino acids marking it as a covalent cell wall anchor common to Gram-positive bacteria. A dedicated extracellular sortase recognizes this anchor, cleaves it, and covalently attaches the protein to a bacterial PG precursor, which is then incorporated into the cell wall (Ton-That *et al*., 2004). Between the internalin domain and the LPXTG cell-surface anchor, the most variable part in this protein family, InlJ bears four *mu*cin *b*inding *p*rotein (MucBP) repeats (Fig. 1-9, Sabet *et al*., 2005).

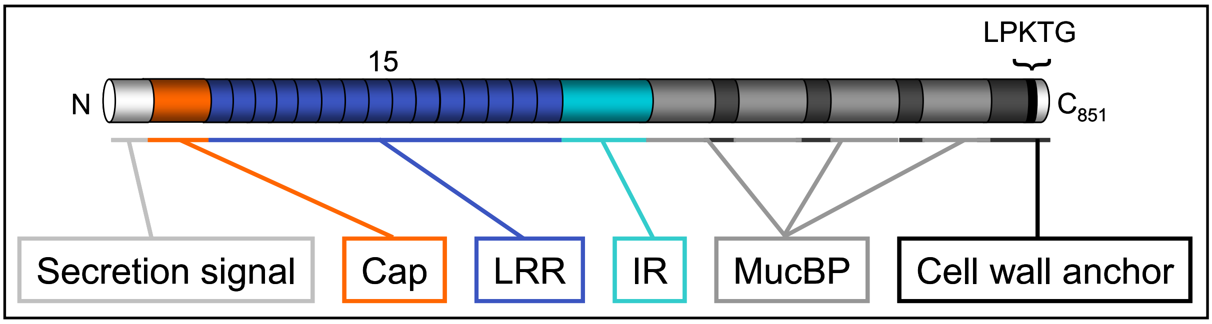


Fig. 1-9: Domain structure of InlJ. The typical internalin domain consisting of cap, LRR and interrepeat (IR) regions is followed C-terminally by four MucBP repeats. The LPKTG-motif at the C-terminus mediates covalent anchoring of InlJ to the cell wall.

An in-frame deletion of *inlJ* significantly reduces bacterial loads: 10-fold in the liver and spleen of intravenously infected wild-type mice and up to 100-fold in the liver, spleen, mesenteric lymph nodes and intestine of orally infected transgenic mice expressing human E-cadherin alongside the murine protein in the intestinal epithelium (Sabet *et al*., 2005). InlJ was correspondingly classified as a novel virulence factor, indispensable for successful infection. A potential receptor has, however, as yet not been identified.

The LRR-consensus sequence of InlJ deviates from that of all other internalin family members, in that each repeat comprises 21 instead of the standard 22 residues. In addition, a hydrophobic residue in one of the LRR-defining positions, immediately following the β-strand, is replaced by a cysteine in InlJ. This results in a novel subtype of the LRR motif. InlJ bears a total of 14 cysteine residues in 15 LRRs – a feature that has not been described for LRR-proteins of known structure.

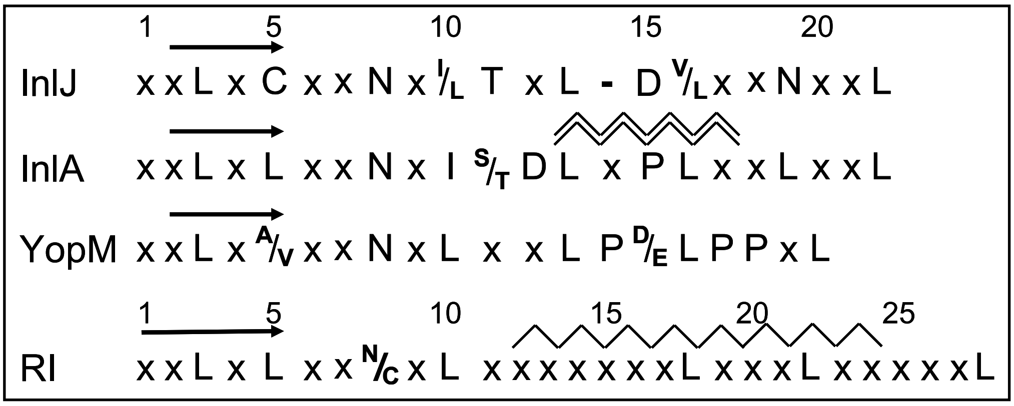


Fig. 1-10: LRR-repeat consensus sequences of InlJ, InlA, YopM (*Yersinia pseudotuberculosis*) and porcine ribonuclease inhibitor (RI). The consensus lengths for InlJ (21 residues), InlA (22), YopM (20) and RI (28/29) differ, as does the pattern of conserved residues. Single-letter residue codes indicate conserved residues, x implies no preference. Arrows indicate β-strands, single zigzag lines α-helices and double zigzag lines 3_{10}-helices.

1.4.2 InlG

In contrast to InlA and InlB, none of the other internalins is known to induce phagocytosis in mammalian cells. Though their explicit function is thus unknown, there is evidence for their importance during infection (Bergmann *et al*., 2002; Ooi *et al*., 2006; Sabet *et al*., 2005).

InlG is the shortest cell wall-bound internalin with five leucine-rich repeats, encoded by the *inlG* gene. In the genome of *L. monocytogenes* EGD-e, *inlG* forms a gene cluster together with *inlH* and *inlE*. The expression of these three genes is not regulated by the general regulator of virulence PrfA. Instead, the proteins are mainly produced under extracellular conditions (Raffelsbauer *et al*., 1998). An *inlGHE*-cluster deletion mutant of *L. monocytogenes* results in significantly reduced bacterial loads in the liver and spleen of orally infected mice (Raffelsbauer *et al*., 1998). Cell-based assays indicate that *inlGHE* cluster has an accessory function in InlA-dependent phagocytosis (Bergmann *et al*., 2002).

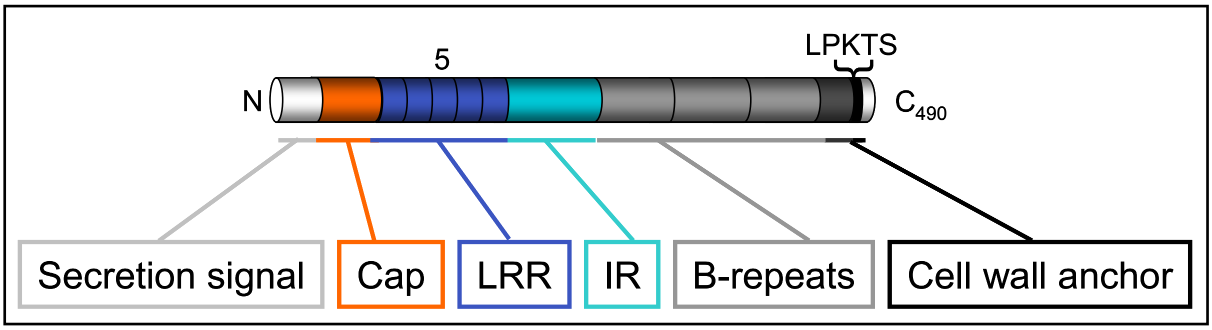


Fig. 1-11: Domain organization of InlG. Five leucine-rich repeats make InlG the shortest of the cell wall-anchored internalins.

2 Aims and Scope

The surface proteome and the secretome of *L. monocytogenes* comprise a multitude of virulence-associated proteins. The identification of novel virulence factors has frequently been achieved by whole-genome comparisons of *L. monocytogenes* with the apathogenic strains. Genes restricted to pathogenic strains are *per se* of interest, as the likelihood of their involvement in bacterial virulence is generally high.

Auto (lmo1076) and the internalins investigated in this study are not found in apathogenic *Listeriae*. This thesis comprises two sections, each of which focuses on a distinct class of proteins. The two groups are, however, united by their involvement in listerial pathogenesis. The aim of this study was to elucidate the contributions of these potential virulence factors to the pathogenicity of *L. monocytogenes*, with a special emphasis on their structural analysis.

2.1 The autolysin Auto

Auto, a listerial autolysin not conserved in *L. innocua*, and recently identified as a novel listerial virulence factor, was to be investigated with respect to its structure and function. For this purpose, the gene was to be cloned to identify stable fragments for recombinant production and purification. To characterize the enzyme mechanism, an assay was required, to allow the assessment of the K_M, substrate range, cleavage site, pH range and inhibition by kinetic and mass spectrometric methods. To correlate these functional characteristics with its structural features, the crystal structure of Auto was to be determined to elucidate the mode of activation and control of Auto during infection.

2.2 Internalin family proteins

Listerial internalins are not only interesting with respect to their involvement in pathogenesis, but also because of their structural properties as leucine-rich repeat proteins. Both these aspects were to be addressed. In an integrated approach, all available internalin genes, encoding InlA, InlB, InlC, InlE, InlF, InlG and InlH, were to be cloned as N- and C-terminal GFP-fusion constructs for heterologous expression in eukaryotic target cells as well as in *E. coli*. The subcellular localization and the influence of GFP-fused internalins in A-431 cells were to be investigated. Furthermore, the crystal structures of InlG and InlJ were to be solved to relate their structural features to possible functions during infection. The geometry of the cysteine-rich InlJ was to be analysed in detail and compared to other LRR proteins, to reveal distinctive structural features of solenoid-shaped repeat proteins, not readily visible by eye.

3 Material and Methods

3.1 Standard Materials

3.1.1 Chemicals, enzymes and kits

If not stated otherwise, all chemicals used in this work were of "*pro analysis*" grade and were purchased from the companies Amersham Biosciences, Difco, Fluka, Hampton Research, Merck, Millipore, QIAGEN, Riedel de Haen, Roche, Roth, Sigma and Stratagene.

Table 3-1: Enzymes.

Enzyme	Source
Benzonase	Merck
TEV-protease	HZI
PreScission® Protease	GE Healthcare
Hen egg-white lysozyme (HEWL)	Fluka
Trypsin	Boehringer
Thermolysin	Boehringer
Subtilisin	Boehringer
Papain	Boehringer
SAP (Shrimp Alkaline Phosphatase)	Roche
T4 DNA Ligase	New England Biolabs
Platinum Pfx DNA Polymerase	Invitrogen
Restriction endonucleases: BamHI, BglII, EcoRI, KpnI, MfeI, NcoI, NdeI, NheI, PstI, XhoI	New England Biolabs

3.1.2 Molecular weight standards

Table 3-2: Molecular weight standards.

Name	Type	Usage	Company
Smart Ladder	DNA	Agarose gel electrophoresis	Eurogentec
Bench Mark Protein Ladder	Protein	SDS-PAGE	Invitrogen
Precision Plus Protein All Blue Standard	Protein	SDS-Page/Western Blot	BioRad

3.1.3 Crystallization screens

The following commercial screens were used to screen for crystallization conditions:

Table 3-3: Crystallization screens.

Screen	Company
The $AmSO_4$	QIAGEN
The Anions	
The Cations	
The Classics I + II	
The Classics Lite	
The Cryos	
The MbClass I + II	
The Sparse Matrix 1-5	
The MPDs	
The Pegs	
The pHClear I + II	
SFP (JCSG+)	
Pro Complex	
PACT	
Grid Screen Ammonium Sulfate	Hampton Research Corp.
Grid Screen PEG/LiCl	
Grid Screen Polyethylene 6000	
Crystal Screen Cryo	
Crystal Screen I + II	
Additive Screen	
Cryo I + II	Jena Bioscience
JBScreen 1 – 10	
JBScreen HTS I	
Wizard I + II	

3.1.4 Bacterial strains

Table 3-4: Bacterial strains.

E. coli **strain**	**Genotype**	**Source**
TOP10	F^-, *mcr* A Δ(*mrr*-*hsd*RMS-*mcr*BC) Φ80 *lac*Z Δ M15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal* U *gal* K *rps*L *end*A1*nup*G	Invitrogen
XL1-Blue super-competent cells	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F′ proAB lacIqZΔM15 Tn10 (*Tet^r*)].*	Stratagene
TG2	*supE hsdΔ5 thiΔ (lac-proAB) Δ(srl-recA) (306 :: Tn10) (*tet^r*) F' [traD36*$proAB^+$ $lacI^q$*lacZΔM15]*	Stratagene
BL21CodonPlus (DE3)-RIL	B F^- *ompT hsdS*(r_B^- m_B^-) *dcm*+ Tet^r *gal* □ (DE3) *endA* Hte *[argU ileY leuW* Cam^r]	Stratagene
BL21 Star	F^- *ompT hsdSB (*rB^-mB^-*) gal dcm rne131 (DE3)*	Invitrogen

3.1.5 Plasmids

The following expression plasmids were used throughout.

pGEX-6p-1 (GE Healthcare)

This vector allows the expression of target protein fused to an N-terminal glutathione S-transferase (GST) tag. A PreScission®-protease cleavage site between tag and target protein allows the fusion protein to be cleaved, releasing the protein of interest. The plasmid confers resistance to ampicillin, allowing a selective pressure to be exerted during bacterial culturing. The vector was used to express the genes $inlG_{36\text{-}460}$ and $inlG_{36\text{-}277}$.

pETM30 (EMBL, Protein Expression Facility)

Genes cloned into the *pETM30* vector are expressed with an N-terminal His_6 and glutathione S-transferase (GST) tag. A TEV-protease cleavage site between tag and target protein allows the fusion protein to be cleaved, releasing the protein of interest. All *aut* constructs as well as $inlJ_{34\text{-}508}$ were cloned into and expressed from this vector.

pETM30GFPfus (HZI)

This vector is based on the *pETM30* plasmid. The presence of the *GFPuv* gene at the 3'-end of the multiple cloning site allows the production of [GST - *target protein* – GFP] fusion proteins. The resulting fusion proteins may be detected by GFP fluorescence or standardized antibody labelling, without the need for target-specific antibodies and are thus ideal for localization assays. As part of the described project, $inlA_{36\text{-}496}$, $inlB_{36\text{-}321}$, $inlE_{36\text{-}342}$, $inlF_{36\text{-}795}$, $inlF_{36\text{-}454}$, $inlG_{36\text{-}460}$, $inlG_{36\text{-}277}$ and $inlH_{36\text{-}519}$ were cloned into this vector.

pEGFP-C1 (BD Biosciences Clontech)

The *pEGFP* vectors can be used for bacterial as well as eukaryotic expression of target proteins with an N-terminal fusion to 'enhanced GFP' (EGPF), a red-shifted variant of wild-type GFP optimized for brighter fluorescence and higher expression in mammalian cells, incorporating the point mutations F64L and S65T. The codon usage in EGFP has been optimized for expression in mammalian cells. The gene constructs $inlA_{36\text{-}496}$, $inlB_{36\text{-}321}$, $inlE_{36\text{-}342}$, $inlF_{36\text{-}795}$, $inlF_{36\text{-}454}$, $inlG_{36\text{-}460}$, $inlG_{36\text{-}277}$ and $inlH_{36\text{-}519}$ were cloned and expressed using this vector.

pEGFP-N3 (BD Biosciences Clontech)

This vector is equivalent to *pEGFP-C1*, with the exception that the GFP-tag is fused to the C-terminus of the target protein. The gene constructs $inlA_{36\text{-}496}$, $inlB_{36\text{-}321}$, $inlE_{36\text{-}342}$, $inlF_{36\text{-}795}$, $inlF_{36\text{-}454}$, $inlG_{36\text{-}460}$, $inlG_{36\text{-}277}$ and $inlH_{36\text{-}519}$ were cloned and expressed using this vector.

3.1.6 Oligonucleotides

Oligonucleotides were used as primers for site-directed mutagenesis and PCR reactions. They were purchased from Invitrogen or Biozym (desalted and HPLC-purified).

Table 3-5: Oligonucleotides for PCR and site-directed mutagenesis. Recognition sites for restriction endonucleases are underlined. Point mutations are marked in red and sites of deletions by underscores.

No.	**Name**	**Length (bp)**	**Sequence (5'→3')**	**T_m (°C)**	**Restriction site**
1	InlA 496 PET-GFP 5'	37	GG AAG CCC ATG GCC GCT ACA ATT ACA CAA GAT ACT CC	64	NcoI
2	InlA 496 PET-GFP 3'	39	CGA CAC GGA TCC CTC GAG TGC CTT AAG TGG CTG CGT CAC	66	BamHI, XhoI
3	InlA 496 eGFP 5'	38	GGA GCG AAT TCA ATG GCT ACA ATT ACA CAA GAT ACT CC	64	EcoRI
4	InlB 321 PET-GFP 5'	34	AC CTT GCC ATG GCC GAG ACT ATC ACC GTG CCA AC	62	NcoI
5	InlB 321 PET-GFP 3'	39	ATA ACT GGA TCC CTC GAG CTC TTT CAG TGG TTG GGT TAC	62	BamHI, XhoI
6	InlB 321 eGFP 5'	37	AC CTT GAA GCT TCA ATG GAG ACT ATC ACC GTG CCA AC	62	HindIII
7	InlE 341 PET-GFP 5'	42	GTA CAA GCC ATG GTG ATT GAA CAC CCA ACC CCT ATT AAC GAA	76	NcoI
8	InlE 341 PET-GFP 3'	45	AAA AAC GGA TCC CTC GAG TTC TGT TAA TGG TTG TGT AAC TGT TCC	74	BamHI, XhoI
9	InlE 341 eGFP 5'	43	G TAC AAG AAT TCA ATG ATT GAA CAC CCA ACC CCT ATT AAC GAA	76	EcoRI
10	InlF PET-GFP 5'	37	GC CCC TCC ATG GCC GCG ACA CAG GTC GCG CCA CCA GC	80	NcoI
11	InlF PET-GFP 3'	45	AC CCG GGA TCC CTC GAG AGT TTT TGG AAG TTT AGC TGT TGC TAG T	76	BamHI, XhoI
12	InlF eGFP 5'	38	G CCC CTG AAT TCA ATG GCG ACA CAG GTC GCG CCA CCA G	76	EcoRI
13	InlF short PET-GFP 3'	45	TTT GTA GGA TCC CTC GAG TTT TAA CGG TTG TGT TAT GAA GCC ACT	74	BamHI, XhoI
14	InlG PET-GFP 5'	39	GG ATG GCC ATG GCC GAG AGT ATT GCG CAA CCA GCG CCA A	78	NcoI
15	InlG PET-GFP 3'	44	CC CGG GCT AGC CTC GAG TGT TTT TGG AAG TTT AGT ACT CGT CGC	76	NheI, XhoI
16	InlG eGFP 5'	41	G GAT GGG AAT TCA ATG GAG AGT ATT GCG CAA CCA GCG CCA A	78	EcoRI
17	InlG eGFP 3'	38	CC CGG GGT ACC TGT TTT TGG AAG TTT AGT ACT CGT CGC	76	KpnI
18	InlG short PET-GFP 3'	44	GTA ATT GCT AGC CTC GAG TAC TAT CGG TTG AAC AAC AGT TCC TG	74	NheI, XhoI
19	InlG short eGFP 3'	38	GTA ATT GGT ACC TAC TAT CGG TTG AAC AAC AGT TCC TG	74	KpnI
20	InlH 519 pET-GFP 5'	34	TG GAG GCC ATG GCC GAG AGC ATC ACG CAG CCA AC	64	NcoI
21	InlH 519 pET-GFP 3'	39	AAG CAT GGA TCC CTC GAG ACC AGC TGT TGG AAG TGT GGT	64	BamHI, XhoI
22	InlH 519 eGFP 5'	35	GTG GAG AAT TCA ATG GAG AGC ATC ACG CAG CCA AC	64	EcoRI

23	Auto 572 pET 5'	34	A CAT GCC ATG GCT GAA ACA ACT AAT GGA GTA GAT	66	NcoI
24	Auto 572 pET 3'	35	CGG GAT CCG ATA TTT AAA TGC TTT TTT GTC CAT CC	66	BamHI
25	Auto 243 pET 3'	36	CGG GAT CCG ATC ATA TTG AGT CAA ATT ATA TGA AGA	66	BamHI
26	Auto 572 QC-FP	31	AA GCA TTT AAA TAT TAA ATC CGA ATT CGA GC	60	
27	Auto 572 QC-RP	31	GC TCG AAT TCG GAT TTA ATA TTT AAA TGC TT	60	
28	Auto 243 QC-FP	29	G ACT CAA TAT GAT TAA ATC CGA ATT CGA G	58	
29	Auto 243 QC-RP	29	C TCG AAT TCG GAT TTA ATC ATA TTG AGT C	58	
30	Auto A83N/A84N FP	57	G CCT CAA ACT AGA AAT AGT TTT TTG AGA AAC AAT AGT ACT CCG ACA TTT CAA CAA AC	73	
31	Auto A83N/A84N RP	57	GT TTG TTG AAA TGT CGG AGT ACT ATT GTT TCT CAA AAA ACT ATT TCT AGT TTG AGG C	73	
32	Auto A83S/A84S FP	57	G CCT CAA ACT AGA AAT AGT TTT TTG AGA TCA TCT AGT ACT CCG ACA TTT CAA CAA AC	74	
33	Auto A83S/A84S RP	57	GT TTG TTG AAA TGT CGG AGT ACT AGA TGA TCT CAA AAA ACT ATT TCT AGT TTG AGG C	74	
34	Auto Deg Forward	36	A CAT GCC ATG GCC GCT AGT ACT CCG ACA TTT CAA CA	66	NcoI
35	Auto Deg Reverse	39	CGG GAT CCG TTA ATC ATA TTG AGT CAA ATT ATA TGA AGA	66	BamHI
36	Auto Loopdel 72-83 FW	40	GAT GCT ATG GCT GCT TTA_GCT AGT ACT CCG ACA TTT CAA C	83	
37	Auto Loopdel 73-83 RV	40	GTT GAA ATG TCG GAG TAC TAG C_TA AAG CAG CCA TAG CAT C	83	
38	Auto E122Q FW	32	GCT CAA GCG GCT CTT CAG AGT AAC TGG GGA AG	80	
39	Auto E122Q RV	32	CT TCC CCA GTT ACT CTG AAG AGC CGC TTG AGC	80	
40	Auto R64D FW	36	C TCG CTT GAA CAA AAT GAC GAT GAT GCT ATG GCT GC	81	
41	Auto R64D RV	36	GC AGC CAT AGC ATC ATC GTC ATT TTG TTC AAG CGA G	81	
42	Auto D194A FW	37	GA AAT GGC CCA AGT TGG GCC TCT AGT TAT TAC AAA GG	79	
43	Auto D194A RV	37	CC TTT GTA ATA ACT AGA GGC CCA ACT TGG GCC ATT TC	79	
44	Auto E61D FW	39	G GAA CCA GTA TTC TCG CTT GAT CAA AAT AGA GAT GAT GC	78	
45	Auto E61D RV	39	GC ATC ATC TCT ATT TTG ATC AAG CGA GAA TAC TGG TTC C	78	
46	Auto K106A K107A FW	47	C ACG CAA GCA ATG GAT TTA TGT GCA GCG TAC AAT TTA TAT CCA TCT G	75	
47	Auto K106A K107A RV	47	C AGA TGG ATA TAA ATT GTA CGC TGC ACA TAA ATC CAT TGC TTG CGT G	75	
48	Auto K132S FW	36	GGA AGA AGT GAA TTA GGG AGC GCC CCT AAT TAC AAC	80	

49	Auto K132S RV	36	GTT GTA ATT AGG GGC GCT CCC TAA TTC ACT TCT TCC	80	
50	Auto K148S FW	40	GGA TCA TAT AAT GGC AGC AGT GTA ACA ATG AAA ACT TGG G	81	
51	Auto K148S RV	40	C CCA AGT TTT CAT TGT TAC ACT GCT GCC ATT ATA TGA TCC	81	
52	Auto K161S FW	40	GG GAA TAT AGT GAT TCC AGC GGT TGG TAT CAA ATC AAC GC	78	
53	Auto K161S RV	40	GC GTT GAT TTG ATA CCA ACC GCT GGA ATC ACT ATA TTC CC	78	
54	Auto 53-243 pET 5'	33	A CAT GCC ATG GAG TCT GAG GAA CCA GTA TTC TC	68	NcoI
55	Auto 53-243 pET 3'	38	AGC GGA TCC TTA ATC ATA TTG AGT CAA ATT ATA TGA AG	70	BamHI
56	InlJ 34-506 Forward	38	AAG TCC ATG GCC ACG ACA GAA ATC ACT ACT GAT AAA GA	70	NcoI
57	InlJ 34-506 Reverse	36	T ATT GGA TCC TTA GGG TTG AGG TGC TTC AAA TGG AG	70	BamHI
58	InlJ Mut FW	42	TT GAA GCA CCT CAA CCC ATA AAG TGA TCC GAA TTC GAG CTC C	84	
59	InlJ Mut RV	42	G GAG CTC GAA TTC GGA TCA CTT TAT GGG TTG AGG TGC TTC AA	84	

3.1.7 Media

Luria Bertani-medium (LB-medium):

Tryptone (Bacto)	10 g/L
Yeast extract (Bacto)	5 g/L
NaCl	5 g/L

SOB-medium:

Tryptone (Bacto)	20 g/L
Yeast extract (Bacto)	5 g/L
NaCl	10mM
KCl	2.5mM

SOC-medium:

SOB-Medium +	
Glucose	2mM
$MgCl_2$	10mM
$MgSO_4$	10mM

BHI-medium (Brain-Heart-Infusion):

Glucose	2mM
$MgCl_2$	10mM
$MgSO_4$	10mM

SWISS-medium:

DMEM	500 mL
FCS	50 mL
Glutamine	2mM
Penicillin (5000 U/mL)/ Streptomycin (5 mg/mL)	5 mL
NEAA solution (Gibco)	5 mL
Pyruvate (100 mM)	5 mL
Glucose	4.5 g/L

3.1.8 Antibiotics

The following antibiotics were used for various cultured cells to ensure the selectivity of liquid and solid media.

Table 3-6: Antibiotics.

Antibiotic	Final concentration
Ampicillin	100 µg/mL
Chloramphenicol	34 µg/mL
Kanamycin	30 µg/mL

3.2 Microbiology

3.2.1 Agar plates

Agar plates for solid-phase cultivation of bacteria were prepared as follows: 16 g agar was added to 1 L LB-medium. The solution was autoclaved and cooled to below 55°C. Antibiotics, as required, were added and the solution was poured into sterile petri dishes (∅10 cm) using a sterile workbench.

Ampicillin-containing plates were stored at 4°C for up to 2 weeks, kanamycin- and chloramphenicol-containing plates for up to 6 weeks. Before use, plates were pre-dried for 5-10 min under the sterile workbench, and bacteria were plated with sterile Drigalski-spatula or inoculation loops.

3.2.2 Liquid culture

Liquid bacterial cultures were grown in Erlenmeyer flasks with 3-4 baffles to ensure sufficient aeration. The flasks with inoculated liquid medium (and antibiotics as required) were incubated at 20 to 37°C and rotation speeds of 160 to 180 rpm, depending on the type of culture: analytical and preparative protein expression, precultures, DNA preparation, cultures for glycerol stocks.

3.2.3 Long-term storage

For long-term storage of bacteria in glycerol stocks, 5 mL LB-medium were inoculated with the clone of interest from liquid culture or agar plate. Antibiotics were added as required. The culture was grown until slightly turbid (37°C and 180 rpm). 900 µL of the culture were mixed gently with 100 µL sterile 87 % glycerol in a 1 mL cryo-tube (Nunc). The mixture was incubated at RT for 10 min and then frozen at –70°C.

For reactivation of frozen cells, a small amount of frozen material was scratched off the surface with an inoculation loop.

3.2.4 Isolation of peptidoglycan from *Listeria innocua*

PG from *L. innocua* strain 34 was isolated essentially as described for *Staphylococcus aureus* PG (Kajimura *et al*., 2005). Briefly, *L. innocua* cells were grown in 1 L BHI-medium o.n. at 37°C and 180 rpm. Cells were harvested by centrifugation (6000 g, 15 min) and washed twice with PBS. The cell pellet was transferred to a cylindrical steel tube and mixed with an approximately equal amount of glass beads (1:1 mixture of diameters 0.11 mm and 1.18 mm). Cell rupture was achieved by elliptical shaking for 30 min in a water-cooled homogenizer (Zellmühle VI4, VIBROGEN). Complete lysis was confirmed microscopically. The lysate was filtered through a glass frit to remove the glass beads, followed by centrifugation for 10 min at 550 g to remove unbroken cells. The insoluble, cell wall containing fraction was pelleted by centrifugation at 30 000 g for 20 min, resuspended in 0.5 % SDS and incubated at 60°C for 30 min for extraction of residual lipids and membrane components. The crude cell wall fraction was pelleted by centrifugation at 30 000 g for 15 min and washed six-fold with water using an Ultra Turrax® mixer (TP 18/10, IKA) to ensure thorough resuspension of the cell wall pellet. Residual cell wall-bound proteins were digested o.n. with 1 mg/mL trypsin in 1 M Tris/HCl pH 7.0 at 37°C. 50 µL of toluene were added to the digest to prevent microbial growth. The preparation was centrifuged (30 000 g, 15 min), washed successively with 1 M Tris/HCl pH 7.0, with 1 M Tris/HCl pH 7.0, 1 M NaCl, with 1 M Tris/HCl pH 7.0 and three times with water. To remove teichoic acids, the preparation was incubated o.n. with 1 mL of 40 % (w/v) aqueous HF at 4°C, followed by extensive rinsing with water to remove residual acid, and finally lyophilized.

3.3 Molecular biology

3.3.1 Preparation of electrocompetent cells

Electrocompetent cells were prepared using *E. coli* TOP10 cells freshly plated on LB-agar. A single colony was picked, transferred to 100 mL of LB-medium, and incubated o.n. (37°C, 180 rpm). 5 mL of this starter culture were transferred to 500 mL LB-medium. Bacteria were grown to an $OD_{600\ nm}$ of 0.4, cooled rapidly and kept on ice for 30 min. Cells were pelleted by centrifugation (4°C, 10 min, 5500 g), washed twice with ice-cold, sterile HEPES buffer (1 mM), once with ice-cold, sterile 10 % (v/v) glycerol, and resuspended in a final volume of 1 mL 10 % (v/v) glycerol (concentration $1\text{-}3 \bullet 10^{10}$ cells/mL). 50 µL aliquots were transferred into pre-cooled Eppendorf tubes and flash-frozen in liquid nitrogen. The electrocompetent cells were stored at -70°C.

3.3.2 Preparation of chemically competent cells

Chemically competent *E. coli* TG2 cells were prepared as follows: a starter culture of 75 mL LB-medium was inoculated from a single colony of a freshly plated agar culture. After o.n. growth at 37°C and 180 rpm, 10 mL of the starter culture was used to inoculate the main culture of 600 mL SOC-medium. The cells were grown at 37°C to an $OD_{600\ nm}$ of 0.5, then cooled and pelleted by centrifugation (4°C, 10 min, 5500 g). The cells were washed and resuspended twice in decreasing volumes of ice-cold and sterile $CaCl_2$ solution (0.1 M). Cold glycerol was then added to a final concentration of 10 % (v/v), and the solution was aliquoted into pre-cooled Eppendorf tubes. The cells were flash-frozen in liquid nitrogen and stored at -70°C.

3.3.3 Transformation of competent bacteria

Competent bacteria were transformed for plasmid amplification or recombinant gene expression. Positive transformants were selected by plating the bacteria onto LB-agar containing antibiotics for which resistance was conferred by the respective plasmid.

Electroporation

1-2 µL of plasmid DNA or ligation mixture were pipetted into a cooled, sterile electroporation cuvette. A slowly thawed 60 µL aliquot of electrocompetent cells was added and the mixture incubated for 1 min on ice. The cuvette was dried and placed into a Gene Pulser^{TM} (BioRad). An electric pulse at 2.5 kV, 200 Ω and 125 µF was applied to the cells, followed by a rapid addition of 1 mL of room-tempered SOC-medium. After incubating at 37°C with gentle shaking for 30-60 min, the transformants were plated on selective LB-agar.

Heat shock transformation

1-2 µL of plasmid DNA, ligation mixture or product of site-directed mutagenesis reactions were added to 50 µL of chemically competent cells (thawed on ice) and incubated for 30 min on ice. A heat shock of 45 (mutagenesis) or 90 s (plasmids or ligations) at 42°C was applied to the mixture, followed by a rapid cooling on ice for 2 min. 800 µL of SOC-medium were added to the cells followed by incubation for 1 h at 37°C with gentle shaking. The transformants were plated on selective LB-agar.

3.3.4 Plasmid preparation

Depending on the amount of plasmid DNA required, 3 (“Mini“), 250 (“Midi“) or 500 (“Maxi“) mL of LB-medium were inoculated with freshly transformed *E. coli* TG2- or TOP10-cells. Bacteria were grown o.n. at 37°C and 180 rpm and the plasmid DNA was isolated with the QIAprep Spin Kit (QIAGEN) or the JETQUICK Plasmid Spin Kit (Genomed). Yields were determined photometrically (see 3.3.5).

3.3.5 Determining DNA concentration and purity

Samples of isolated bacterial plasmid DNA were measured spectrophotometrically to determine the yield and purity. The absorption at 260 nm ($A_{260\ nm}$) was determined, allowing the DNA concentration to be calculated using Eq. 2-1 and assuming that $A_{260\ nm} = 1.0$ is equivalent to 50 µg/mL DNA.

$$c\left[\frac{\mu g}{\mu L}\right] = \frac{A_{260nm} \times D}{20}$$ **Eq. 3-1**

$A_{260\ nm}$ Absorption at 260 nm

c concentration of DNA solution,

D dilution factor

The ratio of $A_{260\ nm}/A_{280\ nm}$ was used to check the sample for protein contamination. Ratios of 1.8 to 2.0 were assumed sufficiently pure to allow subsequent cloning and transformation/transfection experiments.

3.3.6 Agarose gel electrophoresis

DNA samples were analysed by agarose gel electrophoresis using gels prepared from 1 % agarose in TAE, brought to boil in a microwave. After cooling to ~60°C, the agarose solution was supplemented with 5 ‰ ethidium bromide for DNA visualization and poured into an electrophoresis chamber. Plastic combs with teeth of different width were used to create sample pockets of desired volume.

DNA samples were mixed with 10 % loading buffer (10 x) before injection into the sample pockets. Gels were run in TAE buffer with 5 V/cm for 60 min and documented under UV illumination at 254 nm.

TAE buffer:	40	mM	Tris
	20	mM	Sodium acetate
	1	mM	EDTA
	adjusted to pH 8.2 with acetic acid		

10 x DNA loading buffer:	70	% (w/v)	Sucrose
	0.25	% (w/v)	Bromphenol blue
	0.1	M	EDTA

3.3.7 Extraction of DNA from agarose gels

Linearised vectors and PCR products were extracted from agarose gels using the “Gel extraction Kit” (QIAGEN) or the JETQUICK Gel Spin Kit (Genomed).

3.3.8 Digestion of plasmid DNA with restriction endonucleases

Plasmid DNA was digested by distinct combinations of restriction endonucleases (Table 3-1) as required by cloning and control purposes. Buffers and reaction conditions were adapted to the enzyme combinations, according to the New England Biolabs technical reference catalogue.

3.3.9 Dephosphorylation of linearised plasmid DNA

Linearised plasmid DNA was 5'-dephosphorylated prior to ligation to prevent religation of the original vector without the target insert. 25 ng of linearised vector per planned ligation reaction were incubated with 1 U of Shrimp Alkaline Phosphatase at 37°C for 10 min followed by heat inactivation at 65°C for 15 min.

3.3.10 Ligation of DNA fragments

Target genes (inserts) were ligated into linearised, 5'-dephosphorylated vectors using 25 ng linearised plasmid DNA and a 5-fold molar excess of insert. In cases, where this standard protocol was not successful, the molar ratio between insert and vector was varied from 2:1 to 50:1. The mixture was incubated with 1 U of T4 DNA ligase and an ATP-containing ligation buffer at 22°C for 2 h followed by 16 h at 16°C. Typically, 2 µL of ligation reaction mixture were used to transform competent bacteria.

3.3.11 Polymerase chain reaction

Target genes from genomic DNA or template plasmids were amplified by polymerase chain reaction (PCR) to allow (re-)cloning into the respective recipient plasmids. Standard PCR-mixtures contained:

Template-DNA	0.25-0.5	µg
Primer forward	15	pmol
Primer reverse	15	pmol
Deoxyribonucleotides (dNTPs)	300	µM each
Polymerase buffer (10x)	10	%(v/v)
$MgSO_4$	1	mM
Platinum Pfx DNA polymerase	1	U

The reaction was performed in a thermocycler. Typical runs were carried out as follows:

Step 1:	Initial denaturation	2	min	95°C
Step 2:	Denaturation	15	s	95°C
Step 3:	Annealing	30	s	[T_m– 5 to 10 K]
Step 4:	Elongation	3	min	68°C
Step 5:	Final elongation	7	min	68°C

Steps 2-4 were cycled 35 times.

In multiple parallel reactions with primers of different T_m, a temperature gradient was applied during the annealing step and the reactions were carried out in triplicate to ensure optimal yield of amplificate without side products. Amplificates were purified by agarose gel extraction (QIAGEN Gel Extraction kit) or using a PCR purification kit (QIAGEN).

3.3.12 Design and synthesis of deoxyribo-oligonucleotides

Deoxyribo-oligonucleotides were used as primers for PCR and site-directed mutagenesis reactions (see also 3.3.13). To amplify genes of interest for re-cloning, restriction sites compatible with the target vectors were incorporated into tailored overhangs of primers. Cloning strategies were checked using VECTOR NTI (Invitrogen).

Site-directed mutagenesis primers were designed according to the instruction manual of the QuikChange^{TM} Kit (QIAGEN).

3.3.13 Site-directed mutagenesis

Site-directed mutagenesis reactions were performed using the QuikChange^{TM} Kit (QIAGEN). Briefly, desired mutations, insertions or deletions were introduced into a template plasmid during a PCR reaction using an excess of mutagenesis primers. After specific digestion of the methylated, original template DNA, the amplified mutated plasmids were transformed into *E. coli* XL1-Blue cells. Successful mutagenesis was confirmed by DNA sequencing (see 3.3.14) and mass spectrometry analysis of the recombinant variant proteins (see 3.5.6).

3.3.14 DNA sequencing

DNA sequencing reactions were performed by the company GATC. The sequences were determined with the dye terminator method and provided as ABI chromatogram files for download. Chromatograms were analysed and compared using the program CONTIGEXPRESS from VECTORNTI (Invitrogen).

3.4 Protein production and purification

3.4.1 Test expressions

To establish optimal procedures for the recombinant expression of novel gene constructs, small-scale test expressions were performed prior to preparative synthesis. In 50 – 100 mL liquid cultures, parameters such as temperature, expression time and amount of inducing agent were tested. Whole cell protein extracts were then compared using SDS-PAGE (see 3.5.2). Briefly, 5 mL transformed *E. coli* BL21 CodonPlus were incubated at 37°C o.n. and used to inoculate one 100 mL test culture per expression condition. Cultures were grown at 37°C to an $OD_{600\ nm}$ of 0.6 – 0.8, shifted to the desired expression temperature and induced with IPTG. 1 mL samples were taken before induction and at defined points during the test expression. The cells were pelleted, lysed in 100 µL crack-buffer per OD unit and heated at 98°C for 5 min.

Overnight samples were additionally analysed for soluble/insoluble protein ratios. Pelleted cells from 1 mL culture were treated with 100 µL Bug Buster^{TM} (Novagen) per OD unit to extract soluble cytoplasmic proteins. The remaining insoluble proteinaceous material was solubilised in an equal volume of 8 M urea, 50 mM Tris pH 8.0.

Whole cell extracts, soluble and insoluble samples were then analysed by SDS-PAGE.

Crack-buffer:	Tris/HCl pH 6.8	50	mM
	SDS	2	% (w/v)
	Glycerol	10	% (v/v)
	Bromphenol blue	0.1	% (w/v)
	DTT	100	mM

3.4.2 Recombinant protein synthesis

All recombinant proteins in this work were produced in the strain *E. coli* BL21 CodonPlus. Cells carrying the desired plasmid were thawed from a glycerol stock and grown in a 50 mL starter culture in liquid LB-medium supplemented with appropriate antibiotics at 37°C and 180 rpm shaking. After o.n. incubation, the starter culture was used to inoculate 1–4 L of selective LB-medium with a starting OD_{600} of 0.1. This main culture was grown to exponential phase (OD_{600} 0.6–0.8) at 37°C, then shifted to optimal expression temperature as determined by the test expression (20-22°C). Gene expression was induced with 0.1 mM IPTG. Cells were harvested after 2 to 20 h by centrifugation for 15 min at 6000 g.

3.4.3 Production of selenomethionine-labelled protein

The protocol used in this work is mainly based on a previous description (Guerrero *et al.*, 2001). For the production of selenomethionine-labelled pro-Auto A83S/A84S, the expression strain *E. coli* BL21 CodonPlus carrying the plasmid *pETM30-Auto*$_{27\text{-}243}$ *A83S/A84S* was initially grown in LB-medium at 30°C to an $OD_{600\ \mathrm{nm}}$ of 2.7. The cells were washed twice in sterile water and resuspended to an $OD_{600\ \mathrm{nm}}$ of 1.0 in SeMet minimal medium supplemented with 55 mg/L selenomethionine. Expression was immediately induced with 0.1 mM IPTG and continued at 20°C o.n.. Purification was performed as described for the native protein (see 3.4.4 - 3.4.10). Quantitative incorporation of selenomethionine was confirmed by mass spectrometry (see 3.5.6).

Minimal medium (MM):

Basic medium:	NH_4Cl	1	g
	KH_2PO_4	3	g
	$Na_2HPO_4 \cdot 7\ H_2O$	6	g
	add H_2O	1	L
	autoclaved		
Supplementary solution:	Glucose•H_2O	22	g
	$MgSO_4 \cdot 7\ H_2O$	0.61	g
	Thiamine•HCl	11.2	mg
	$Fe_2(SO_4)_3 \cdot H_2O$	10.4	mg
	add H_2O	100	mL
	sterile filtered and added to basic medium.		

3.4.4 Cell lysis

All steps were performed on ice or at 4°C to prevent proteolytic degradation of the target proteins. The cell pellet was resuspended in 15 mL lysis buffer per 1 L culture.

Lysis buffer:	50	mL	PBS
	5	µL	Benzonase® (25 U/µL)
	1	tablet	Complete Mini protease inhibitor cocktail (Roche)
PBS (pH 7.4):	0.14	M	NaCl
	3	mM	KCl
	10	mM	Na_2HPO_4
	2	mM	KH_2PO_4

The endonuclease Benzonase® in the lysis buffer degrades DNA and RNA in the lysate, reducing the viscosity of the sample and prevents the target protein from being precipitated by unspecific complex formation with nucleic acids.

Cell were disrupted by French Press (SLM Aminco) at a pressure of 1000-1500 psi (6.9 – 10.3 MPa, three runs) or with a high pressure cell disruption device (Constant Systems, 20 kPa, two runs). The lysate was controlled visually for loss of opacity and change of colour. Soluble and insoluble components were separated by centrifugation at 37 000 g for 90 min. The target protein was isolated from the supernatant containing the soluble cytoplasmic proteins. The insoluble fraction, containing cell debris and insoluble proteins was resuspended in 8 M urea. Target protein production and solubility was analysed by SDS-PAGE and Coomassie Blue staining, using 2 µL samples of each the soluble and the insoluble fraction.

3.4.5 Affinity chromatography

Affinity chromatography is based on the strong and specific binding properties of a so-called affinity tag to an immobilized ligand. Via terminal fusion of the affinity tag to the target protein, it can be specifically purified out of a bacterial cell lysate.

The 25-kDa enzyme glutathione S-transferase (GST) binds its natural substrate glutathione with high affinity. It is, therefore, used as an affinity tag fused to the target protein of interest. In this work InlG', InlJ', and all Auto constructs were produced as N-terminal GST-fusion proteins.

Recombinant fusion proteins were enriched from soluble cell extracts on glutathione-coated sepharose 4B beads (Amersham Biosciences). After coupling with PBS-equilibrated beads for 1 h at 4°C, the beads were washed extensively with PBS. The flow-through was checked for proteinaceous material with a flow-through UV monitor at 280 nm. After washing, the fusion protein was either competitively eluted with buffer containing 10 mM free reduced glutathione (InlG'), or proteolytically cleaved off the resin-bound GST with TEV-protease (all *pETM30* constructs, see also 3.4.6).

3.4.6 Proteolytic cleavage of GST-fusion proteins

Since GST has a considerable size (25 kDa) and is furthermore homodimeric, it was removed from the target protein before crystallization setups. The expression vector *pETM30* used for the production of N-terminal GST fusion proteins encodes a TEV-protease cleavage site between the GST-tag and the target protein. Fusion proteins coupled to glutathione sepharose beads were washed in TEV digest buffer and then incubated with 250 µg TEV-protease for 18 h. Free target protein was eluted from the column in a final volume of max. 100 mL.

TEV digest buffer:	Tris/HCl pH 8.0	50	mM
	EDTA	0.5	mM
	DTT	1	mM

In the case of GST-InlG', expressed from the vector *pGEX-6p-1*, the fusion protein was eluted from the glutathione-sepharose beads and the GST-tag was removed with PreScission® Protease (140 U), after adjustment of the eluate to optimal cleavage conditions.

Elution buffer:	Tris/HCl pH 8.2	50	mM
	NaCl	150	mM
	DTT	1	mM
	reduced glutathione	10	mM

PreScission® digest buffer:	Tris/HCl pH 7.0	50	mM
	EDTA	1	mM
	DTT	1	mM

After cleavage, the sample was extensively dialysed against glutathione-free buffer, and GST as well as GST-tagged PreScission® protease removed by a second affinity chromatography step using fresh glutathione-sepharose.

3.4.7 Dialysis

For buffer exchanges and desalting procedures, protein samples were transferred into a dialysis device appropriate to the sample volume and molecule size, and dialysed o.n. or longer against an at least 100-fold excess of liquid at 4°C.

3.4.8 Ion exchange chromatography

By ion exchange chromatography (IEC), protein mixtures were separated depending on their electrostatic properties. Samples were applied to chromatography columns with a high capacity to bind charged molecules. Target protein binding to the column was ensured by the buffer used for the mobile phase, as the net-charge of a given protein is strongly dependent on the pH of the solvent. In this work, the strong anion exchanger MonoQ (GE Healthcare) was used in the column formats HR 5/5 and HR 10/10. After binding and washing, the bound proteins were eluted in a NaCl gradient.

The following table lists the buffer pairs used for IEC of the proteins purified within the scope of this work.

Table 3-7: IEC buffers for proteins purified in this work.

Target protein	Low-salt buffer	High-salt buffer
pro-Auto	10 mM Tris/HCl pH 8.0, 1 mM DTT	10 mM Tris/HCl pH 8.0, 1 mM DTT, 1 M NaCl
InlJ'	10 mM Tris/HCl pH 8.0, 5 mM DTT, 50 mM NaCl	10 mM Tris/HCl pH 8.0, 5 mM DTT, 1 M NaCl
InlG'	10 mM HEPES, pH 7.5, 5 mM DTT	10 mM HEPES, pH 7.5, 5 mM DTT, 1 M NaCl

3.4.9 Gel permeation chromatography

Gel permeation chromatography (GPC) utilizes the correlation of molecular retention times in a porous gel matrix and the physical size of the molecules. The gel matrix contains pores of a given range of diameters, causing a negative correlation between the size of a molecule and its retention time on the column. GPC experiments were performed on ÄKTA FPLC Systems (GE Healthcare) with HiLoad Superdex 75, and 200 (formats 10/30, 16/60 and 26-60) and Superose-12 (10/30 and 16/50) columns. For the mobile phase, low salt IEC buffers were used for the different proteins (see Table 3-7).

3.4.10 Concentration of protein solutions

Protein solutions were adjusted to desired concentrations by ultracentrifugation using Vivaspin 2, -6, and -20 concentrators with appropriate MW cut-offs. The pore size was chosen to be at least 10 kDa smaller than the target protein. The absorption at 280 nm was checked against the flow-through until the desired protein concentration was reached (see 3.5.1).

3.5 Protein biochemical methods

3.5.1 Photometric quantification of protein concentration

Concentrations of purified protein solutions were determined via the measurement of $A_{280\ \mathrm{nm}}$ against buffer. The molar extinction coefficient ε_{280} for each protein was calculated *in silico* with the program VECTORNTI (Invitrogen) or the PROTPARAMTOOL on the EXPASY server (www.expasy.org).

According to the Beer-Lambert law:

$$\mathrm{A} = \varepsilon \cdot \mathrm{c} \cdot d \qquad \textbf{Eq. 3-2}$$

the concentration c can be calculated if the values for A, d, and ε are known.
Since this method cannot distinguish between different proteins in a mixture, it was only used when the target protein was expected to account for the main protein fraction in the solution, with negligible contaminations.

3.5.2 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteinaceous samples were analysed by SDS-PAGE: samples of 2 – 10 µL volume were mixed with 2 µL of SDS-containing 8 x sample buffer and brought to a final volume of 12 µL with water. The samples were incubated for 5 min at 96°C to ensure complete unfolding. After a short centrifugation, the samples were applied to the stacking gel with a Hamilton syringe. Gels were run at 40 mA per gel for 25-35 min, depending on the polyacrylamide percentage and the expected range of MW of the proteins present in the samples.

Table 3-8: Composition of SDS-polyacrylamide gels.

Solution	**Separating gel**		**Stacking gel**
	12 %	15 %	5 %
Acrylamide/Bisacrylamide 30 % (w/v)/0.8 % (w/v)	12 mL	10 mL	1.5 mL
4 x Lower buffer	7.6 mL	5 mL	
4 x Upper buffer			2.5 mL
10 % SDS	0.3 mL	0.2 mL	
H_2O	10 mL	4.7 mL	5.9 mL
TEMED	40 µL	20 µL	15 µL
25 % APS	100 µL	50 µL	25 µL

8 x Sample buffer:

10 % SDS	16	mL
Glycerol	4	mL
Tris/HCl, pH 6.8	2.2	mL
β-Mercaptoethanol	800	µL
Bromphenol blue	1	spatula tip

4 x Upper buffer:	Tris/HCl, pH 6.8	0.5	M
	SDS	0.4	% (w/v)
4 x Lower buffer:	Tris/HCl, pH 8.8	1.5	M
10 x SDS-PAGE running buffer:	Tris Base	30.2	g
	Glycine	144.1	g
	SDS	10	g
	add H_2O	1	L
Staining solution:	Ethanol	30	% (v/v)
	Acetic acid	10	% (v/v)
	Coomassie R-250	0.25	% (w/v)
Destaining solution:	Ethanol	40	% (v/v)
	Acetic acid	10	% (v/v)
Dehydrating solution:	Methanol	50	% (v/v)
	Glycerol	3	% (v/v)
	H_2O	47	% (v/v)

3.5.3 Transfer of proteins to membranes (Western Blot)

Protein samples in SDS polyacrylamide gels were transferred to and immobilized on polyvinylidene difluoride (PVDF) membranes for subsequent immunodetection or N-terminal sequencing experiments. Freshly run gels were equilibrated for 15 min in transfer buffer, together with 2 pieces of gel-sized Whatman paper. The PVDF-membrane was activated in 100 % methanol for 1 min and briefly equilibrated in transfer buffer. For the transfer, the gel was placed onto the membrane, and both placed between two layers of soaked Whatman paper onto the anode of a semi-dry blot apparatus. The blot was run with 15 V for 20 min. Success of the transfer was checked via reversible staining of the membrane with Ponceau red. Samples intended for N-terminal sequencing were cut out of Ponceau red stained membrane bands.

Transfer buffer:	CAPS	10 %
	Methanol	10 %
	H_2O	80 %
Ponceau red staining solution:	Ponceau S	0.2 %
	Acetic acid	2 %

3.5.4 Immunodetection of immobilized proteins

Immobilized target proteins were detected using a specific antibody directed against the GST-tag of the fusion protein (α-GST-HRP conjugate, Amersham Biosciences).

After the transfer (see 3.5.3), the membrane was ‘blocked’ to prevent non-specific binding of the antibody by incubation for 1 h with 50 mL of 5% (w/v) skim milk in PBST (PBS supplemented with 0.1% (v/v) TWEEN-20) at 4°C and washed 5 x 5 min in PBST. The saturated membrane was then incubated with the antibody for at least 1 h at RT or at 4°C o.n.. After antibody-binding, the membrane was washed again 3 x 10 min in PBST. The peroxidase-bound antibody was visualized on the membrane using an enhanced chemiluminescence (ECL) kit (Amersham Biosciences).

For documentation, protein bands on PVDF membranes were stained using H_2O_2 and 4-chloro-1-naphtol. The membrane was washed 2-3 times for 10 min first in water and then in PBS. The staining solution was then applied to the membrane for up to 15 min. The reaction was terminated by a further washing step in water.

Staining solution:	PBS	47	mL
	0.3 % 4-chloro-1-naphtol in methanol	3	mL
	30 % H_2O_2	25	µL

3.5.5 N-terminal sequencing

Protein samples were identified and evaluated by N-terminal sequencing. After separation by SDS-PAGE and transfer to a PVDF membrane, Ponceau red stained protein bands were excised for analysis. All sequencing experiments were carried out by Rita Getzlaff (HZI, Braunschweig) as described (Edman & Begg, 1967).

3.5.6 Mass spectrometry

Both protein and PG fragment (muropeptide) samples were analysed by mass spectrometry (MS), by Dr. Manfred Nimtz and Undine Felgenträger (HZI, Braunschweig) as follows:

Proteins

Recombinantly produced, purified polypeptides were checked by MALDI-TOF (*Matrix Assisted Laser Desorption Ionisation – Time Of Flight*) or ESI-MS (*Electrospray Ionisation*) methods for purity and intactness of the expected amino- and carboxy-terminal ends. Furthermore, overall mass shifts were determined in order to confirm site-directed amino acid exchanges and to quantify incorporation of selenomethionine. Protein samples were dialysed against water or 5 mM ammonium acetate prior to MS-analysis.

Muropeptides

0.5 mg purified PG from *L. innocua* were dissolved in 500 µL buffer and digested o.n. at 37°C with 10 µg Auto_{act}. Undigested PG was removed by centrifugation, and the released soluble muropeptides from the supernatant were desalted and analysed by MS as follows:

1. Desalting

PG preparations were purified on reversed-phased C_{18} ZipTip pipette tips (Millipore Corporation) previously washed with an aqueous solution containing 0.5 % HCOOH and 65 % MeOH and equilibrated with a solution containing 0.5 % HCOOH and 3 % MeOH. After washing (0.5 % HCOOH, 3 % MeOH), the purified and concentrated PG fragments were eluted in 4–5 µL 1.0 % HCOOH, 65 % MeOH.

2. ESI-MS/(MS)

The PG samples obtained by digest with Auto_{act} were dissolved in a 1:1-mixture of $\text{MeOH:H}_2\text{O}$ to a concentration of approximately 3 pmol/µL, and filled into gold-coated nanospray glass capillaries (Protana). The tip of the capillary was placed orthogonally in front of the entrance hole of a QTOF II mass spectrometer (Micromass) equipped with a nanospray ion source, and a voltage of approximately 1000 V was applied. For collision-induced dissociation (CID) experiments, parent ions were selectively transmitted from the quadrupole mass analyser into the collision cell. Argon was used as the collision gas with a kinetic energy setting from 25 to 60 eV for optimal fragmentation. Resulting daughter ions were separated by the orthogonal time-of-flight (TOF) mass analyser. Recorded spectra were enhanced using the MAX.ENT.3 software package (Micromass).

3. Reduction of PG samples

Approximately 10 µg of PG fragments dissolved in 50 µL water were mixed with the same volume of NaBH_4-solution (2 mg/mL in 1 % NH_3) and incubated for 24 h at RT. The reaction was quenched by addition of 10 µL acetic acid. After evaporation of solvents in a speed vac concentrator, boric acid was removed by addition of 100 µL of 1 % acetic acid in MeOH followed by evaporation of solvents in a stream of nitrogen (repeated 5-fold).

3.5.7 Dynamic light scattering

Dynamic light scattering (DLS) is a powerful method to probe the properties of molecules in solution (Clark *et al.*, 1970). Monochromatic and coherent laser light passing through an aqueous sample of suspended particles is scattered in all directions, producing a random pattern of varying intensity. Since the molecules are in random thermal (Brownian) motion, the scattering changes continuously, leading to changing interference of the scattered waves and a signal of fluctuating intensity. Analysis of the time-dependence of this intensity fluctuation may be used to determine the diffusion coefficient of the molecules in the sample and, in conjunction with details about the viscosity, allows the hydrodynamic radius to be calculated. Macromolecular (including protein) solutions are said to be monodisperse if all solute molecules in solution adopt identical states of aggregation. The monodispersity of a solution such as of a protein is known to correlate with the likelihood of its successful crystallization (Thibault *et al*., 1992). Multimers or aggregates of different sizes represent a severe obstacle for successful crystallization.

DLS-measurements were performed for every purified protein solution to determine the percentage of polydispersity, to confirm the expected MW and to establish the state of oligomerization. If aggregates and monomers could not be sharply separated by GPC, individual chromatography fractions were analysed in order to determine the optimal limits for pooling.

3.5.8 Limited proteolysis

Both pro-Auto and its variant Δ27-51/Δ72-83 were analysed by limited proteolysis using the endoproteases trypsin, chymotrypsin, thermolysin, subtilisin and papain in a protease/enzyme ratio of 1:100. Digests were incubated at RT in 100 mM Tris/HCl pH 8, 10 mM $CaCl_2$ (in case of papain plus a trace of L-cysteine). 8 µL samples were taken after 5, 15, 30, 60 and 120 minutes, and the reaction was stopped by the adding AEBSF (for trypsin, chymotrypsin, subtilisin), phosphoramidon (thermolysin), or E-64 (papain). Then, the samples were supplemented with SDS sample buffer, heated and analysed by SDS-PAGE. Stable cleavage products were N-terminally sequenced.

3.5.9 Renaturing SDS-PAGE

Autolytic activity was tested by renaturing SDS-PAGE with 15 % polyacrylamide gels containing 0.4 % KCN-extracted and autoclaved *L. monocytogenes* EGD-18 (kindly provided by Reiner Munder, HZI Braunschweig) or lyophilized *Micrococcus lysodeiktikus* (Sigma, ATCC No. 4698) cells based on previous descriptions (Cibik & Chapot-Chartier, 2004; McLaughlan & Foster, 1997). Enzyme samples were prepared in conventional SDS sample buffer without boiling, and the gel was run with 40 mA for 30 min. The gel was washed in water for 15 min, then transferred to renaturing buffer (100 mM sodium acetate, pH 4.8) and incubated at 37°C for several minutes. Lytic bands appearing as clear zones in the opaque gel were counterstained with 0.1 % (w/v) methylene blue in 0.01 % (w/v) KOH.

Multiple-substrate zymography was performed likewise, but using polyacrylamide gels with columns containing different substrates. The multi-gels were produced by sequentially polymerizing 5 juxtaposed vertical gel columns, each containing 20 mg autoclaved cells from *Streptococcus infantis* (DSM12492), *Nocardioides kribbensis* (DSM16314), *Nocardioides aromaticivorans* (DSM15131), *Microlunatus phosphovorus* (DSM10555), *Anaerococcus lactolyticus* (DSM7456), *Brachybacterium nesterenkovii* (DSM9573), *Micrococcus luteus* (DSM20030), *Microbacterium lacticum* (DSM20427), *Listeria monocytogenes* EGD-18 and lyophilized *Micrococcus lysodeiktikus* (Sigma, ATCC No. 4698).

3.5.10 Photometric enzyme kinetics

The autolytic activities of pro-Auto and Auto_{act} were compared by determining the decrease of $OD_{600\ nm}$ of a suspension of log-phase *L. innocua* cells in 100 mM Tris/HCl pH 8. Optimal pH, as well as V_0- and K_M-values were assessed in a turbidity assay using *Micrococcus lysodeiktikus* cells as described for HEWL (Stenesh, 1984). Briefly, lyophilized *M. lysodeiktikus* cells were dissolved in a final volume of 300 µL 100 mM Tris/HCl pH 4.8, 50 mM NaCl and mixed with 50 µL Auto_{act} solution (final concentration 33 nM). The $OD_{450\ nm}$ was measured at intervals of 10 s over 100 s. Autolytic activity of Auto_{act} and the variants E122Q, and Δ27-51/Δ72-83 were compared at a substrate concentration of 0.2 mg/mL. The (photometer-dependent) limitation of measurable substrate concentrations was expanded by applying an experimentally determined correlation function to transform non-linear, high OD values (1-1.8) into reliable substrate concentrations.

To assess the enzyme activity over the acidic pH-range, different buffer systems were used for the assay: glycine/HCl for pH 3.0–3.6, acetate buffer for pH 3.6–5.8.

3.5.11 Accessible thiol group quantification

The quantitative determination of accessible thiol groups was performed using Ellman’s reagent (5,5’-Dithio-bis(2-nitrobenzoic acid), DTNB) as previously described (Ellman, 1958; Riddles *et al.*, 1983). Briefly, InlJ' (1.4 µM) was incubated with 1 mM DTNB (Sigma). The release of TNB (2-nitro-5-mercaptobenzoic acid) upon reaction of DTNB with free thiols was measured photometrically at 412 nm after 20 min of reaction time. 1 to 10 µM L-cysteine solutions were used to establish a standard curve. All samples were prepared using the same

DTT-free buffer for optimal accuracy. The assay was performed under native conditions in 10 mM Tris/HCl pH 8.0, 50 mM NaCl and repeated under denaturing conditions in 6 M guanidinium hydrochloride pH 8.0.

3.6 Protein crystallization

Proteins were crystallized by hanging- and sitting-drop vapour diffusion methods.

Fast screening

Each protein solution was initially screened for lead crystallization conditions using commercial screens (QIAGEN, see also section 3.1.2) in 96-well sitting-drop plates. 200-600 nL drops composed of equal amounts of protein- and crystallization solution were pipetted using a MOSQUITO robot (TTP LabTech Ltd.). The plates were sealed with Manco^{TM} Crystal Clear tape (Jena Bioscience) and incubated at different temperatures (25°C, 20°C, 12°C and 4°C).

Optimization

Initial hit conditions were optimized manually in 24-well hanging drop and sitting drop formats, using drop volumes of 2-3 µL by varying the physico-chemical parameters such as precipitant and protein concentration, pH, ionic strength and temperature to produce crystals suitable for X-ray diffraction experiments.

3.7 X-ray data collection

Protein crystals larger than 30-50 µm in the smallest dimension were harvested from their mother liquor and flash-cooled in liquid nitrogen. 30 % Glycerol (pro-Auto A83S/A84S, pro-Auto Δ27-51/Δ72-83, InlJ') and 30 % PEG 4000 (InlG') were used as cryoprotectants.

X-ray diffraction data were collected on a rotating copper anode generator and an R-AXIS IV^{++} image plate detector (Rigaku) and at synchrotron beamlines BW7A, X12 and X11 (EMBL, DESY, Hamburg) and BL1 (BESSY, Berlin) on charge-coupled device detectors (MarCCD, Marresearch).

3.8 Data processing and structure solution

X-ray diffraction data were processed with Denzo and Scalepack of the HKL-Suite (Otwinowski & Minor, 1997). Data were further processed using the CCP4 Suite of software (Collaborative Computational Project 4, 1994).

3.8.1 Pro-Auto A83S/A84S and Δ27-51/Δ72-83

The phases for pro-Auto variant A83S/A84S were obtained using multiple anomalous dispersion (MAD) data from selenomethionine derivatized protein crystals (see 3.4.3). The substructure was solved and phases calculated using the programs SHELXC, -D and -E through the graphical user interface HKL2MAP (Pape & Schneider, 2004; Schneider & Sheldrick, 2002; Sheldrick, 2002; Sheldrick, 2004), and AutoSolve (PHENIX, Adams *et al.*, 2004). The AutoBuild function of PHENIX was used for automated model building. The initial partial model was completed manually with COOT (Emsley & Cowtan, 2004) and refined with CNS (Brünger *et al.*, 1998), REFMAC5 (Murshudov *et al.*, 1997) and SHELXL (Sheldrick & Schneider, 1997). Twinning tests were done with Xtriage (Zwart *et al.*, 2005).
The structure of pro-Auto Δ27-51/Δ72-83 was solved by molecular replacement with PHASER (McCoy *et al.*, 2005; Storoni *et al.*, 2004) using the MAD-derived model and refined with REFMAC5 using restrained refinement and TLS protocols.

3.8.2 InlJ'

The structure of InlJ' was solved by molecular replacement. A search model encompassing LRRs 1-13 of InlA (PDB code 1O6V) was generated by CHAINSAW (Schwarzenbacher *et al.*, 2004) and positioned by maximum-likelihood based molecular replacement (PHASER, McCoy *et al.*, 2005; Storoni *et al.*, 2004). The interrepeat domain model of InlB (PDB code 1H6T) was located in a second round. The partial models were subjected to rigid body refinement, simulated annealing combined with individual B-factor refinement and energy minimization (CNS, Brünger *et al.*, 1998), followed by several cycles of manual rebuilding (COOT, Emsley & Cowtan, 2004) and restrained refinement including TLS-protocols (REFMAC5, Murshudov *et al.*, 1997).

3.8.3 InlG'

A molecular replacement search model was generated manually, based on the structure of InlB, taking on the cap domain and LRRs 1-5, and fusing them directly to the IR domain. The amino acid sequence of the model was automatically adapted to $\text{InlG}_{36\text{-}277}$ with BRAGI (Reichelt *et al.*, 2005). The molecular replacement solution obtained from PHASER (McCoy *et al.*, 2005; Storoni *et al.*, 2004) was subsequently refined with CNS (Brünger *et al.*, 1998) and REFMAC5 (Murshudov *et al.*, 1997). Manual rebuilding of the molecular model was done with COOT (Emsley & Cowtan, 2004).

Quality of all final models was evaluated with PROCHECK (Laskowski *et al.*, 1993) and MOLPROBITY (Davis *et al.*, 2007).

3.9 Cell culture

3.9.1 Transfection of A-431 cells

The human epithelial carcinoma cell line A-431 (ATCC No. CRL-1555) was cultivated in SWISS-medium, containing Dulbecco's modified Eagle's medium (DMEM, Invitrogen) supplemented with 10 % foetal calf serum (FCS, Sigma, lot 34K3396) and 4.5 g/L glucose at 37°C, 7 % CO_2 and 100 % relative humidity.

For transfection and fluorescence experiments, the cells were grown on gelatine-coated coverslips in 24-well plates. Plasmids encoding *inl-GFP* fusion genes were transfected using the reagent FUGENE (Boehringer) as follows: 24 µL DMEM were mixed with 1.2 µL FUGENE, 0.4 µg of plasmid DNA added and the mixture incubated for 20 min at RT. The cells were provided with 0.5 mL fresh DMEM with FCS, and the transfection mixture was slowly pipetted onto the cells. After careful mixture, the transfection plate was incubated o.n. in the incubator. Transfected cells were examined microscopically for their overall appearance and subcellular localization of GFP fluorescence.

3.9.2 Fluorescence staining of the actin cytoskeleton

The cells were fixed on coverslips by adding 0.5 mL prewarmed (37°C) 4 % *para*-formaldehyde in PBS for 20 min. The coverslips were washed with PBS and permeabilized with the detergent 0.2 % Triton-X-100 (in PBS) for 90 s. After removing the detergent and washing with PBS, the coverslips inverted (cells facing downward) and placed onto 20 µL-drops of phalloidin Alexa Fluor 594 (Molecular Probes) solution (2 U/ml), in 5 % horse serum (PAA, Pasching) and 1 % BSA in PBS, followed by incubation in a moist chamber for 1 h. The cells were washed three times in PBS and fixed by embedding them in warmed Mowiol wax (containing 4 % n-propyl gallate as bleaching protectant) on glass slides. Samples were examined on an inverted microscope (Axiovert 153TV, Zeiss) equipped for epifluorescence, using a 63 x apochromatic objective with a numeric aperture of 1.4, adapted to chromatic aberration. Digital images were acquired using a cooled CCD-camera (Princeton Research Instruments), controlled by IPLab Spectrum P software (Scanalytics). Images were analysed and edited using IPLab Spectrum P and Adobe Photoshop 5.0.

4 Results

4.1 The autolysin Auto

In this study, the soluble domain of Auto (residues 27-243), encompassing the N-terminal part of previously unknown function and the cell wall hydrolase domain, was investigated. The protein was produced recombinantly in *E. coli* and purified chromatographically. It was characterized enzymatically using various PG substrates and enzyme variants. Its crystal structure was solved and refined at 2.55 Å resolution.

4.1.1 Cloning

Two Auto constructs were created, allowing the recombinant production of full-length mature Auto (residues 27-572, only lacking the N-terminal secretion signal) and a C-terminally truncated form, terminating after the catalytic hydrolase domain (residues 27-243). DNA coding for Auto$_{27\text{-}243}$ and Auto$_{27\text{-}572}$, respectively, was amplified by PCR from genomic DNA of *L. monocytogenes* EGD-18 using the oligonucleotide pairs 23/25 and 23/24 (see Table 3-5). After ligation into the recipient vector, stop-codons were introduced with the mutagenesis primer pairs 28/29 for Auto$_{27\text{-}243}$ and 26/27 for Auto$_{27\text{-}572}$.

4.1.2 Protein production and purification

Mature, full-length Auto$_{27\text{-}572,}$ including the four C-terminal cell wall attaching GW-modules, was poorly expressed in *E. coli* BL21 CodonPlus and approximately half of the produced protein was lost to insoluble inclusion bodies. Auto$_{27\text{-}243}$, lacking the GW-modules, by contrast, was strongly overproduced and was largely soluble. For this reason, only the truncated Auto$_{27\text{-}243}$ was analysed further. It was produced as an N-terminal GST fusion protein. The GST-tag was removed by TEV-protease cleavage on the affinity chromatography column, such that three amino acids Gly-Ala-Met were retained at the N-terminus of the recombinant protein. Key biophysical data on the GST fusion protein and the free enzyme after cleavage are summarized in Table 4-1.

Table 4-1: Physico-chemical parameters of GST-Auto$_{27\text{-}243}$, Auto$_{27\text{-}243}$ and the major degradation product Auto$_{84\text{-}243}$, as calculated with VectorNTI (Invitrogen).

Parameter	**GST-Auto$_{27\text{-}243}$**	**Auto$_{27\text{-}243}$**	**Auto$_{84\text{-}243}$**
Length (aa)	469	220	160
Molecular weight (kDa)	53.6	24.8	18.2
Molar extinction coefficient	100310	52180	52180
Isoelectric point (pI)	5.95	5.2	9.25

Affinity chromatography was followed by anion exchange and gel permeation chromatography. During first purification trials, Auto$_{27\text{-}243}$ was found to be particularly sensitive to degradation, resulting in a stable fragment of 18 kDa (Fig. 4-1, Table 4-1).

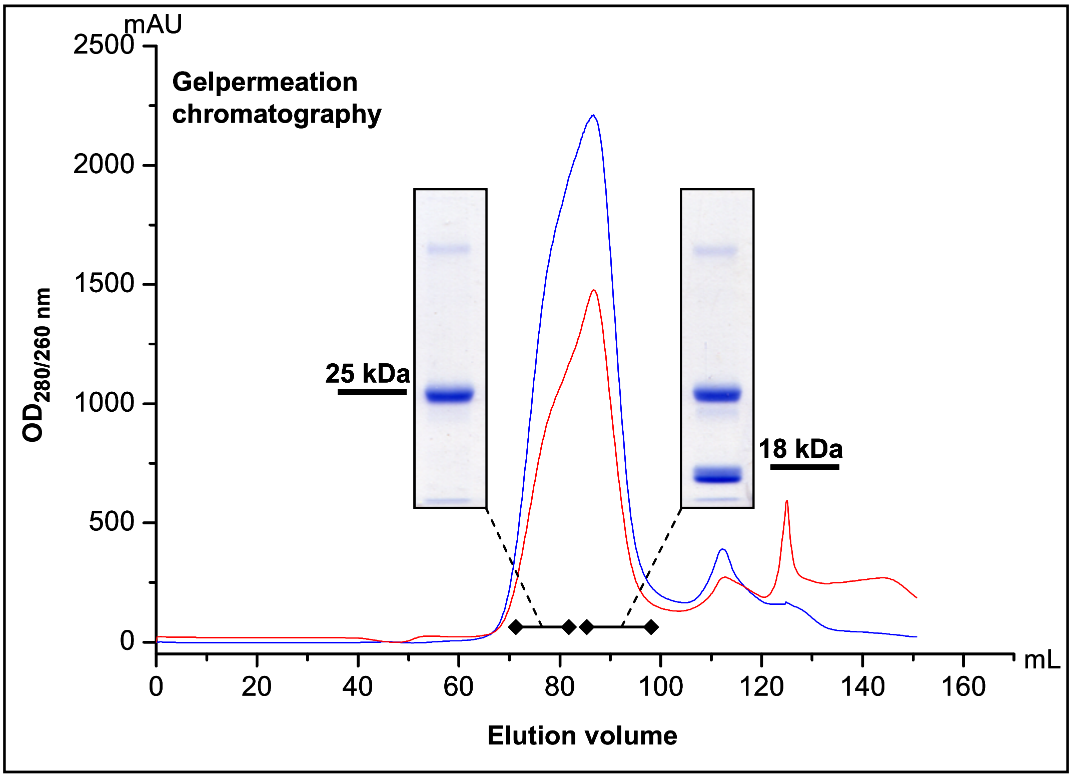


Fig. 4-1: Degradation of Auto$_{27\text{-}243}$ during purification. Blue: $OD_{280\text{ nm}}$, red: $OD_{260\text{ nm}}$. The GPC peak is asymmetric, of which the right half consists of a mixture of intact protein and degradation product.

N-terminal sequencing revealed the protein to be cleaved mainly after Ala83 and to a lesser extent after Ala84 and Ser86, leading to the loss of the N-terminal region of unassigned function. The positive net charge of the degradation product at pH 8.0 allowed for its efficient separation from the intact protein by anion exchange chromatography (see Table 4-1).

GPC was omitted in subsequent preparations because the protein was already pure and crystallisable after the IEC, and further prolongation of the purification procedure would have enhanced its degradation.

4.1.3 The soluble domain of Auto is an inactive pro-enzyme

The autolytic activity of both $\text{Auto}_{27\text{-}243}$ and $\text{Auto}_{84\text{-}243}$ was investigated in a renaturing SDS-PAGE assay using autoclaved *L. monocytogenes* cells as substrate. Only the degradation product $\text{Auto}_{84\text{-}243}$ was observed to be active, whereas the full-length protein $\text{Auto}_{27\text{-}243}$ had no hydrolytic activity (Fig. 4-2A). To confirm that this result was not due to unsuccessful refolding of the uncleaved protein within the gel, the assay was repeated in solution, monitoring the decrease of turbidity in a buffered solution of *log*-phase *L. innocua* cells (Fig. 4-2B).

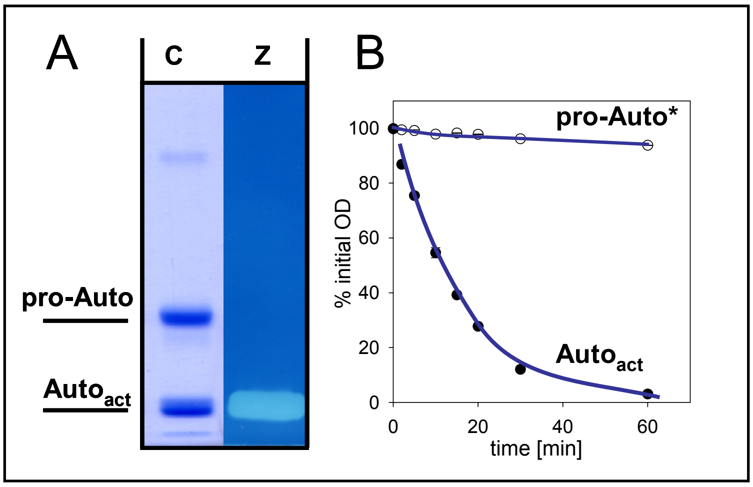


Fig. 4-2: The 18 kDa degradation product $\text{Auto}_{84\text{-}243}$ (Auto_{act}) of $\text{Auto}_{27\text{-}243}$ (pro-Auto) is autolytically active. (A) Coomassie stained (C) sample compared to zymogram (Z). Both pro-Auto and Auto_{act} are present in the sample, but only Auto_{act} generates a lytic band in the zymogram. **(B)** Turbidity assay with log-phase *L. innocua* cells. Pro-Auto lacks activity, whereas Auto_{act} leads to clearing of the cell suspension by substrate cell lysis. *This plot was obtained from pro-Auto variant A83S/A84S, which could be purified with the least Auto_{act} contamination.

Again, only the 18 kDa fragment lacking the N-terminal region is able to degrade bacterial cell walls. This indicates that the N-terminal amino acids 27-83 exert an autoinhibitory function. Accordingly, the inactive 25 kDa pro-enzyme $\text{Auto}_{27\text{-}243}$ will, henceforth, be referred to as pro-Auto, the active 18 kDa processed form $\text{Auto}_{84\text{-}243}$ as Auto_{act}. In addition to the discovery of the autoinhibitory function of the N-terminal 57 residues, this experiment also demonstrates that the cell wall binding GW-modules are not required for the hydrolytic activity of Auto$_{act}$, a property distinguishing it from phage-derived listerial endolysins (Loessner *et al.*, 2002).

4.1.4 Auto is activated by unspecific proteolytic cleavage

How is the "spontaneous" activation of pro-Auto brought about? To inhibit potentially specific proteases, the major cleavage site between Ala83 and Ala84 was modified by generating the pro-Auto variants A83S/A84S and A83N/A84N. Both variants are, however, no more stable than the wild type protein. Instead, N-terminal sequencing of the products indicates new cleavage sites after Arg82 for A83S/A84S and after Asn84 for A83N/A84N. Activation is thus to some extent unspecific, though restricted to an "*activation region*" comprising residues 82-86. Limited proteolysis experiments reveal that chymotrypsin and subtilisin cause extensive degradation of pro-Auto (Fig. 4-3) and consequently loss of enzymatic activity. Trypsin, thermolysin and papain, by contrast, all cleave and activate pro-Auto without, however, causing further degradation. Auto$_{act}$ was stable even after o.n. exposure to these proteases.

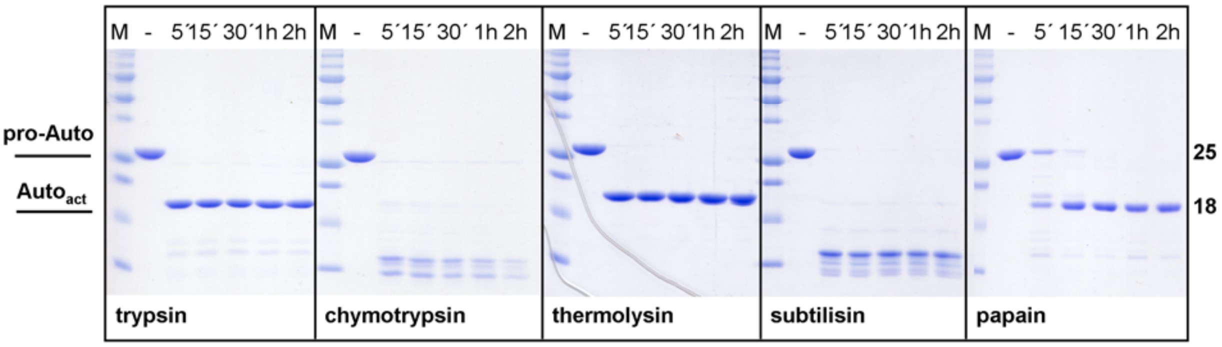


Fig. 4-3: Limited proteolysis of pro-Auto. Trypsin, thermolysin and papain produce the 18 kDa Auto$_{act}$ fragment. M: Marker, -: undigested pro-Auto.

Removing the activation region in the variant Δ27-51/Δ72-83 results in a protein that is no longer activated by proteolytic cleavage (see below) and is not degraded by the proteases listed above, with the only exception of subtilisin, which again largely degrades the protein (Fig. 4-4).

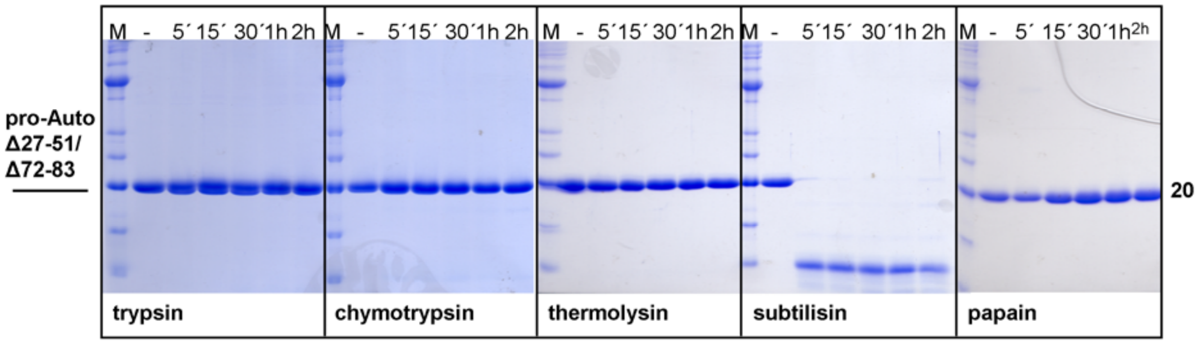


Fig. 4-4: Limited proteolysis of pro-Auto Δ27-51/Δ72-83. Trypsin, chymotrypsin, thermolysin and papain do not cleave the protein. M: Marker, -: undigested pro-Auto Δ27-51/Δ72-83.

4.1.5 Auto$_{act}$ is active at acidic pH

To investigate the enzymatic properties of Auto$_{act}$ in more detail, a standardized photometric lysozyme activity assay was used, based on monitoring the decrease of turbidity in a solution of lyophilized *M. lysodeikticus* cells (Stenesh, 1984). Auto$_{act}$, similar to lysozyme, deviates from classical Michaelis-Menten kinetics due to moderate substrate inhibition, indicated by a maximum in its velocity profile (Fig. 4-5A). The K_M (0.16 mg/ml) and K_S (2.4 mg/ml) were accordingly estimated using a model that involves non-productive enzyme-substrate complexes, with two substrate chains binding to different subsites within the substrate binding site.

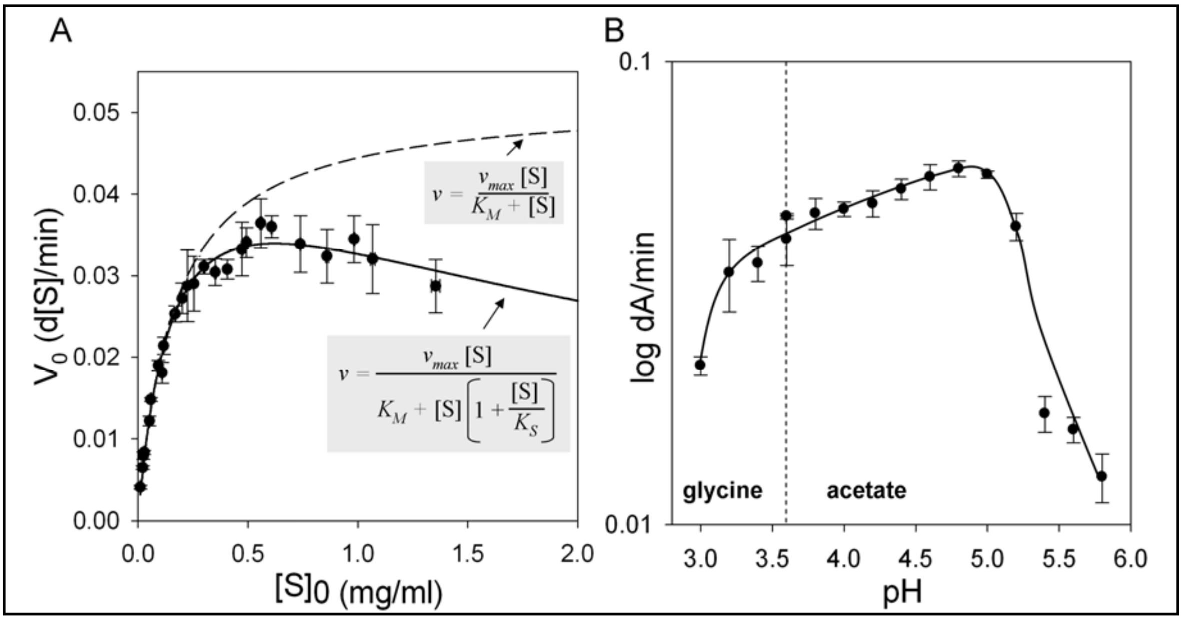


Fig. 4-5: Effect of substrate concentration and pH on the activity of Auto$_{act}$. (A) V_0 plotted against substrate concentration. Auto$_{act}$ displays moderate substrate inhibition. Data were fitted to the indicated equations. K_M = 0.16 mg/ml; K_S = 2.4 mg/ml. Dashed line: theoretical curve without substrate inhibition. **(B)** V_0 against pH. Auto$_{act}$ is active in acidic environments, with a pH-optimum of 4.8. The dashed line indicates the change of the buffer system from glycine to acetate.

Auto_{act} is active in a pH-range between 3.2 and 5.2 with a pH-optimum of 4.8. Below and above this range, activity decreases dramatically (Fig. 4-5B), suggesting a mode of controlling the enzyme in a physiological context.

Like other autolysins associated with the negatively charged cell wall, Auto is a basic protein, the full length secreted protein $\text{Auto}_{27\text{-}572}$ having a predicted pI of 9.6. Despite a reasonable scepticism towards theoretically calculated pI values, it is interesting to note, that the charged amino acids are not evenly distributed throughout the sequence. Instead, a pronounced negative charge cluster is located within the N-terminal region. Eleven acidic amino acids between residues 28 and 55 lead to a theoretical overall acidic pI of 5.2 for pro-Auto that, however, switches to 9.3 upon activation. In the native, full-length protein $\text{Auto}_{27\text{-}572}$, the positive charge of the four C-terminal GW-modules (theoretical pI 10) diminishes the effect of the acidic cluster such that the pI merely shifts from 9.6 to 9.8 upon activation.

4.1.6 Substrate specificity

Sequence-based annotations identify Auto as a muramidase, glucosaminidase or *N*-acetylmuramyl-L-alanine amidase. To probe the substrate range of Auto_{act} in more detail, its lytic activity on different PG-types was tested by multi-substrate renaturing SDS-PAGE (Fig. 4-6). Fig. 4-7 schematically depicts the tested PG-types.

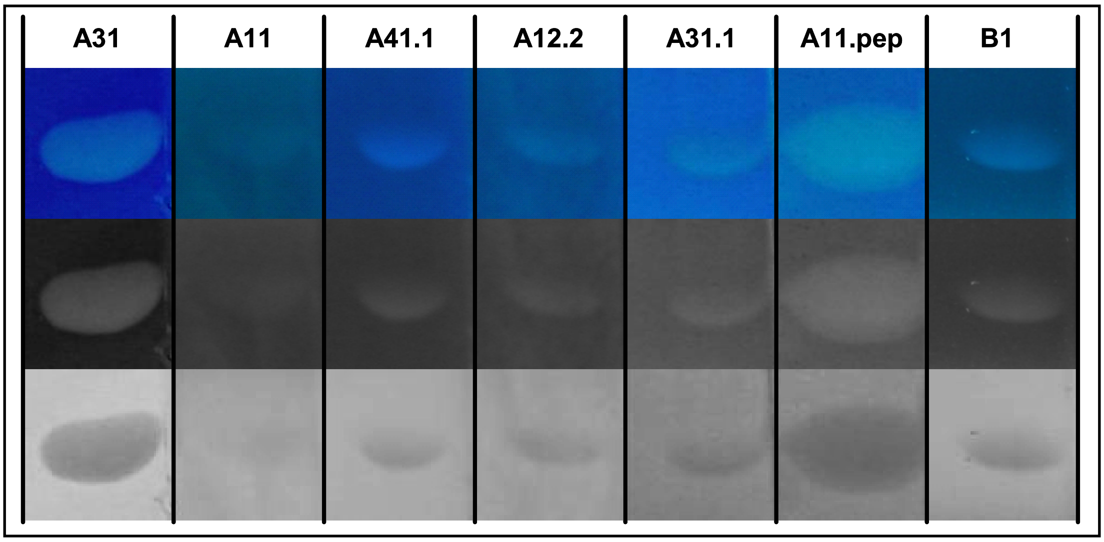


Fig. 4-6: Multi-substrate zymogram of Auto_{act}. All seven tested PG-types are hydrolysed, as seen by the lytic bands. Upper row: contrast-enhanced (Adobe Photoshop) original images of methylene blue counterstained gels, middle row: Grey-scaled images, lower row: inverted grey-scaled images.

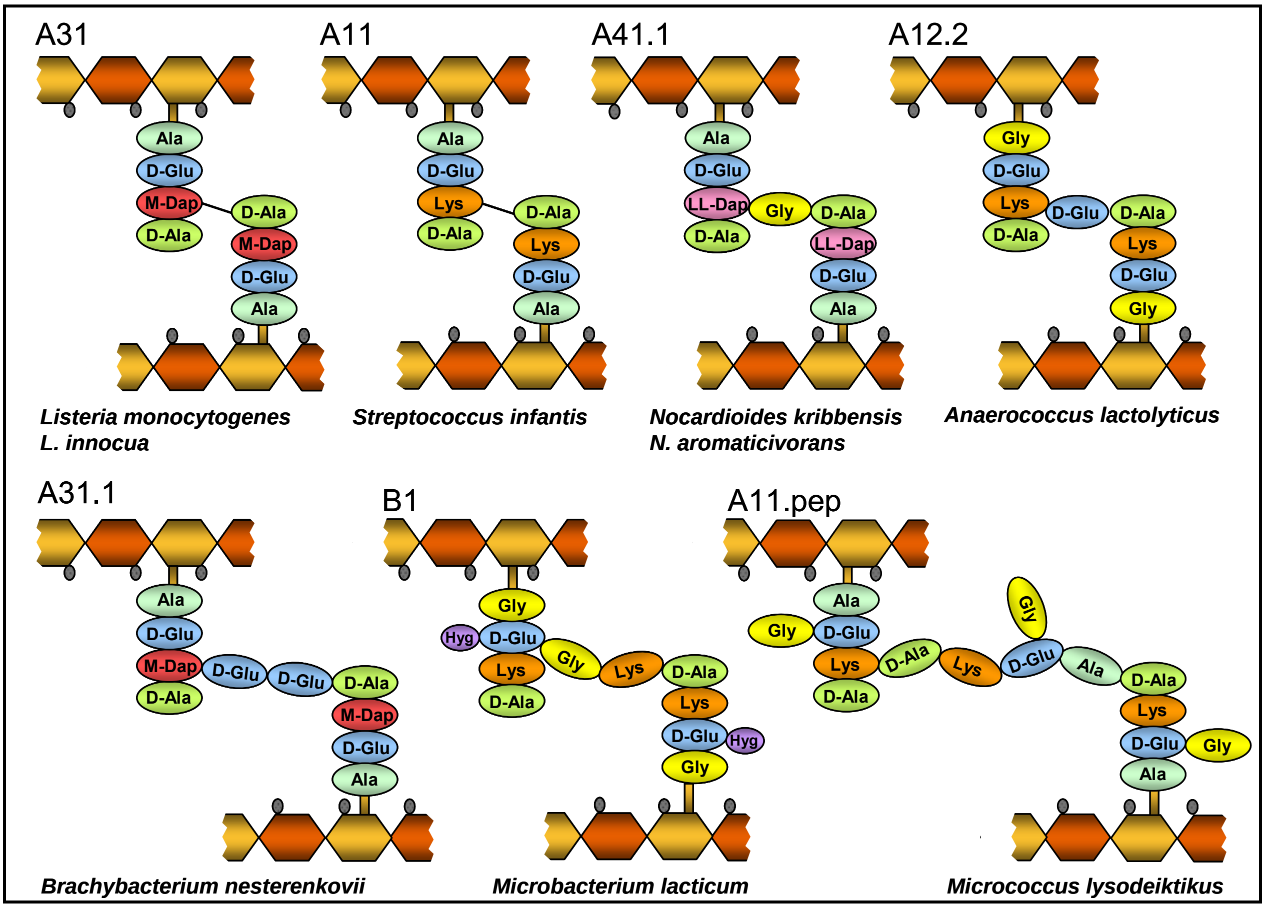


Fig. 4-7: Peptidoglycan types analysed with respect to their susceptibility to Auto$_{act}$ hydrolysis. Note that the glycan strands are invariant whereas the residue sequence and composition of bridging peptides vary substantially. Orange- and red-coloured hexagons: NAM and NAG respectively, grey ellipses: acetyl moieties, orange bars: lactyl moieties of NAM, coloured ellipses: peptide bridge residues, amino acids labelled in three-letter-code, *m*DAP: *meso*-diaminopimelic acid, LL-DAP: L,L-diaminopimelic acid, Hyg: threo-3-hydroxy-glutamic acid.

All substrates are digested by Auto$_{act}$, irrespective of their peptide bridge composition (Fig. 4-6). Auto thus cleaves the invariant carbohydrate backbone, making it an *N*-acetylmuramidase, an *N*-acetylglucosaminidase or a lytic transglycosylase. The fact that different cell-wall types are cleaved at different speeds is likely to be due to differences in the detailed chemical composition of the PG-substrates. Modifications that could affect the accessibility of the scissile glycosidic bond include the degree of crosslinking, *N*- and *O*-acetylation states and incorporation of other components such as (lipo)teichoic acids with or without additional glycosylation.

4.1.7 Auto is an *N*-acetylglucosaminidase

To unambiguously determine the specificity of Auto, soluble muropeptides derived from the Auto$_{act}$ digestion of purified *L. innocua* PG were analysed by mass spectrometry. The smallest cleavage product identified corresponds to a crosslinked 'dimer' of disaccharide-tripeptide and disaccharide-tetrapeptide [NAG-NAM-Ala-Glu-*m*Dap-Ala-*m*Dap-Glu-Ala-NAM-NAG] with a mass of 1788.7 Da (Fig. 4-8), confirming that Auto$_{act}$ cleaves the substrate within the *N*-acetylglycan backbone. This 'dimer' was also detected in an *N*-deacetylated form (mass shift –42) as well as in one- (+42) and two-fold (+48) *O*-acetylated form. In addition, larger, incompletely digested products linked through their peptide bridges were also detected. Interestingly, the major fractions of the incompletely digested fragments were *N*-deacetylated, implying that Auto cleaves *N*-deacetylated PG less efficiently than its acetylated form. This has previously been observed for chicken egg-white lysozyme (Amano *et al.*, 1977; Kamisango *et al.*, 1982). The PG samples were, furthermore, found to contain amidated *m*DAP. Again this confirms similar previous observations for *L. monocytogenes* (Boneca *et al.*, 2007) and *B. subtilis* (Horsburgh *et al.*, 2003).

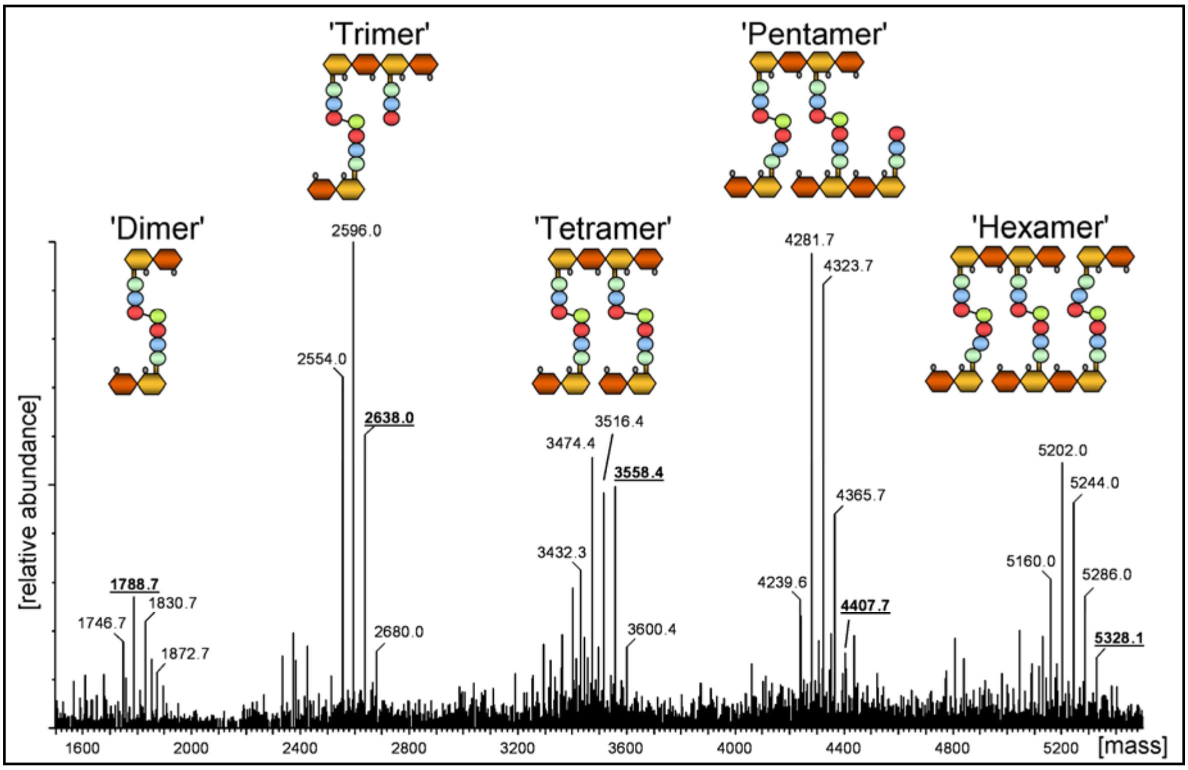


Fig. 4-8: Mass spectrum of Auto$_{act}$-derived muropeptides. Muropeptides corresponding to underlined mass peaks are depicted as cartoons. Neighbouring peaks are shifted by +42 (*O*-acetylated) or –42 (*N*-deacetylated). Glycosidic bond breaks in crosslinked 'tetra-' 'penta-' and 'hexamers' were deduced from mass shifts of +18 (incorporation of water) and from mass shifts of +2 per cleaved bond in reduced samples (not shown). Orange- and red-coloured hexagons: NAM and NAG respectively, grey ellipses: acetyl moieties, orange bars: lactyl moieties, coloured circles: peptide bridge residues (mint: L-Ala, blue: D-isoGlu, red: *m*DAP, green: D-Ala).

To identify the anomeric C1 atom released upon cleavage by Auto$_{act}$, digestion products were reduced to the respective alcohols using $NaBH_4$ and compared to untreated digestion products by MS/MS analysis. A shift in mass of +2 Da indicates the incorporation of two hydrogen atoms at the anomeric C1 atom of free NAM or NAG. Comparing reduced and non-reduced muropeptide 'dimers' (Fig. 4-9) indicates that $NaBH_4$-treated samples contain reduced NAG, indicating that Auto$_{act}$ cleaves the NAG-(1,4)-NAM bond. Auto is thus an *N*-acetylglucosaminidase and not a muramidase or amidase as predicted. *O*-acetylated fragments (mass shift +42) are not present in the reduced samples due to the release of ester-linked acetate moieties during $NaBH_4$-reduction. A detailed listing of the detected mass peaks in both non-reduced and reduced spectra is provided in Table 4-2.

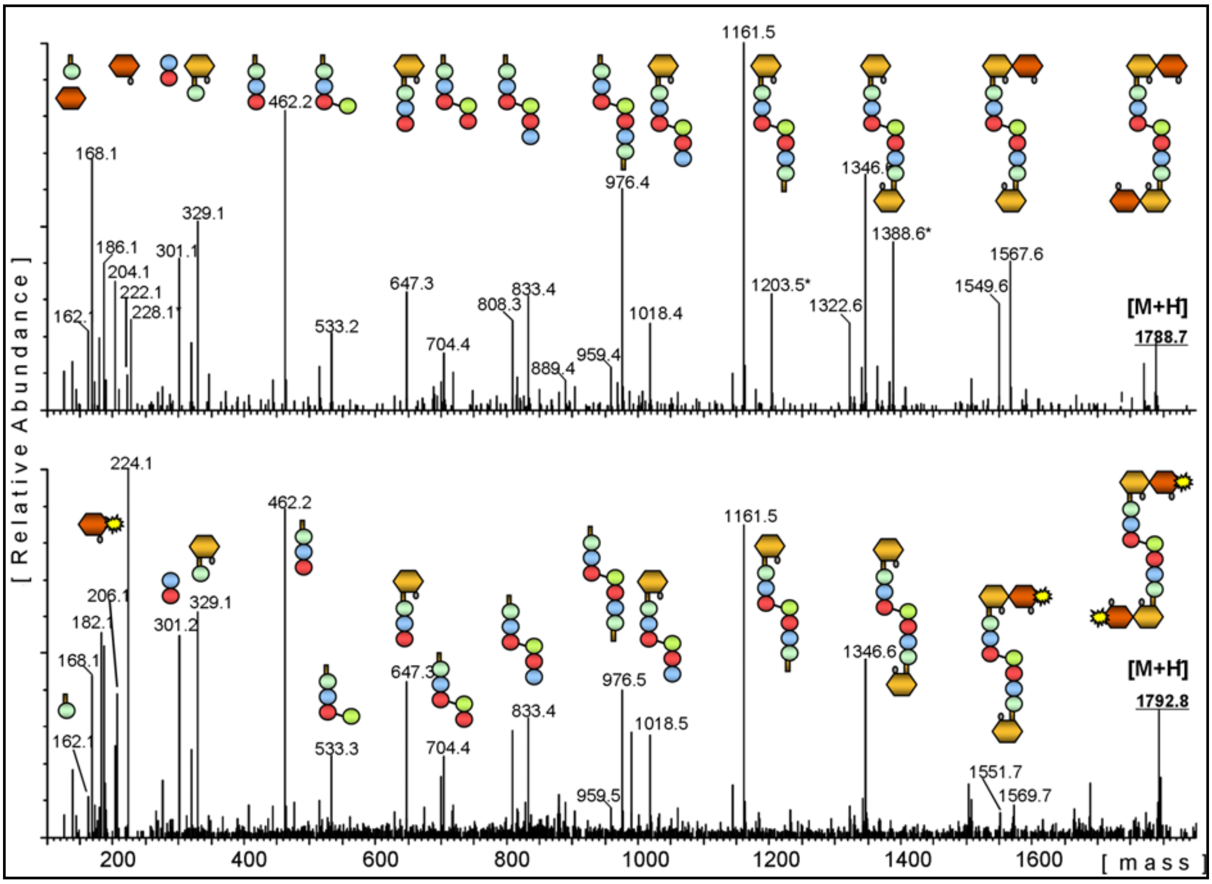


Fig. 4-9: Collision induced dissociation (CID) analysis of a muropeptide 'dimer'. Upper panel: non-reduced sample, lower panel: $NaBH_4$-reduced sample. Reduction of an anomeric C1-atom leads to a mass shift of +2. Orange- and red-coloured hexagons: NAM and NAG respectively, grey ellipses: acetyl moieties, orange bars: lactyl moieties, coloured circles: peptide bridge residues (mint: L-Ala, blue: D-*iso*Glu, red: *m*DAP, green: D-Ala). Reduced C1-atoms are marked with a yellow asterisk. Mass peaks marked with a black asterisk correspond to a shift of +42, indicating *O*-acetylation.

Table 4-2: Muropeptide ‘dimer’ daughter ions detected in a comparative CID experiment. Reduced NAG residues are shifted by [M+H]+2. As anticipated, *O*-acetylated fragment ions are absent in the reduced sample (n.d. = not detected), *, **, *** = one-, two-, threefold dehydrated ion; *O*Ac = *O*-acetyl; G = glucosamine.

Non-reduced		**Reduced**	
Muropeptide	**[M+H]$^+$**	**Muropeptide**	**[M+H]$^+$**
Lac-Ala / G	162.1	"	162.1
NAG***	168.1	"	168.1
NAG*	204.1	NAG$^{*\ red}$	206.1
NAG	222.1	NAG^{red}	224.1
***O*Ac**-NAG***	228.1	**n.d.**	
Glu-Dap^{amid}	301.1	"	301.1
NAM**-Ala	329.1	"	329.1
Lac-Ala-*i*Glu-Dap^{amid}	462.2	"	462.2
Lac-Ala-*i*Glu-Dap^{amid}-Ala	533.2	"	533.2
NAM*-Ala-*i*Glu-Dap^{amid}	647.3	"	647.3
Lac-Ala-*i*Glu-Dap^{amid}-Ala-Dap^{amid}	704.4	"	704.4
G-NAM*-Ala-*i*Glu-Dap^{amid}	808.3	"	808.3
Lac-Ala-*i*Glu-Dap^{amid}-Ala-Dap^{amid}-*i*Glu	833.4	"	833.4
NAM*-Ala-*i*Glu-Dap^{amid}-Ala-Dap^{amid}	889.4	"	889.4
Lac-Ala-*i*Glu-Dap^{amid}-Ala-Dap^{amid}-*i*Glu-Ala-Lac	976.4	"	976.4
NAM*-Ala-*i*Glu-Dap^{amid}-Ala-Dap^{amid}-*i*Glu	1018.4	"	1018.4
NAM*-Ala-*i*Glu-Dap^{amid}-Ala-Dap^{amid}-*i*Glu-Ala-Lac	1161.5	"	1161.5
***O*Ac**-NAM*-Ala-*i*Glu-Dap^{amid}-Ala-Dap^{amid}-*i*Glu-Ala-Lac	1203.5	**n.d.**	
NAM-Ala-*i*Glu-Dap^{amid}-Ala-Dap^{amid}-*i*Glu-Ala-NAM	1346.6	"	1346.6
***O*Ac**-NAM-Ala-*i*Glu-Dap^{amid}-Ala-Dap^{amid}-*i*Glu-Ala-NAM	1388.6	**n.d.**	
NAG*-NAM-Ala-*i*Glu-Dap^{amid}-Ala-Dap^{amid}-*i*Glu-Ala-NAM	1549.6	**NAG**$^{*\ red}$ - "	1551.7
NAG-NAM-Ala-*i*Glu-Dap^{amid}-Ala-Dap^{amid}-*i*Glu-Ala-NAM	1567.6	**NAG**^{red} - "	1569.7
NAG-NAM-Ala-*i*Glu-Dap^{amid}-Ala-Dap^{amid}-*i*Glu-Ala-NAM-**NAG**	1788.7	**NAG**^{red}- " -**NAG**^{red}	1792.8

4.1.8 Crystal structure determination of the pro-Auto variant A83S/A84S

Wild-type pro-Auto and the cleavage site variant A83S/A84S crystallized as tetragonal bi-pyramids (see inset in Fig. 4-13). The crystals grew only when using 17–45 % ethanol as precipitant, without any additional components. Crystals of A83S/A84S were generally more robust than those of wild-type protein and diffracted X-rays to 2.3 Å resolution. Four MAD data sets of SeMet-derivatized A83S/A84S were collected at beamline X12 (EMBL, DESY, Hamburg) and the data were processed in space group $P4_32_12$. After MAD-phasing, automated and manual model building into the experimental electron density (PHENIX, COOT) a complete model of the catalytic domain could be built (Fig. 4-10).

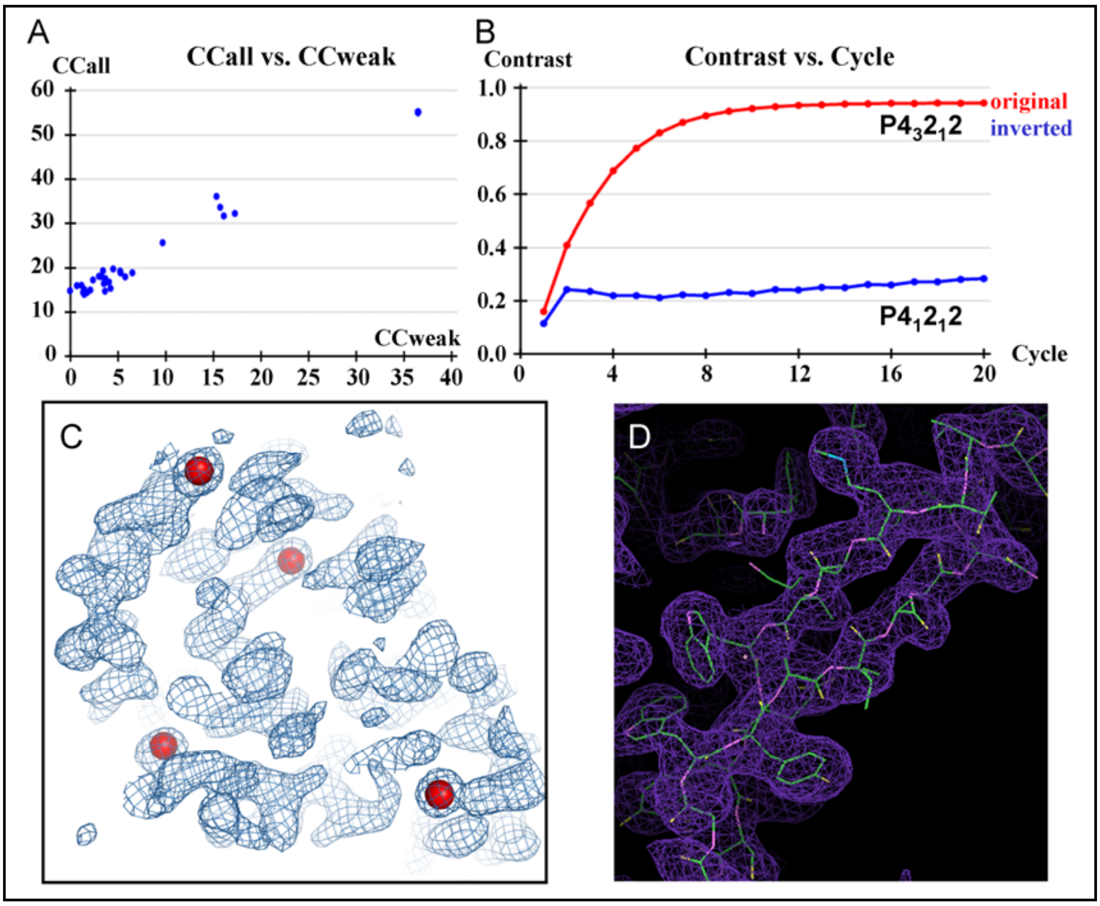


Fig. 4-10: Experimental MAD-phasing: Selenium substructure solution, density modification and initial model building. (A) Se substructure solution with SHELXD. A clearly separate solution (in the upper right corner of the plot) of high correlation coefficients between observed and calculated heavy-atom structure factors indicates the Se substructure to be correctly solved. **(B)** Density modification with SHELXE. The contrast of the electron density map for both possible heavy-atom enantiomorphs is plotted versus the cycles of density modification. The higher contrast reveals the space group $P4_32_12$ to be correct. **(C)** Initial 'experimental' electron density ($|Fo|/\alpha_{exp}$, blue mesh, contoured at 1σ). Selenium atoms (red spheres) are located in regions interpretable as selenomethionine side chains. **(D)** Automated model building with AutoBuild (PHENIX). Parts of the amino acid sequence could be automatically traced into the initial electron density.

This model lacked only two disordered regions, namely the N-terminal residues 27-55 and a loop between residues 75 and 88, including the previously determined activation region.

Non-crystallographic symmetry

Calculating a self-rotation function with diffraction data from pro-Auto A83S/A84S, previous to structure solution, did not reveal any non-crystallographic symmetry (NCS) axes between the monomers in the asymmetric unit (Fig. 4-11A). However, subsequent structure solution revealed that two of the three monomers in the asymmetric unit are related by a non-crystallographic two-fold rotation axis in addition to the crystallographic symmetry axes. This NCS axis is masked in the self-rotation function, because it is inclined by only ~13° relative to the crystallographic *c*-axis (Fig. 4-11B). By being perpendicular to a crystallographic two-fold axis, the non-crystallographic symmetry (NCS)-axis generates a second, equivalent NCS-axis perpendicular to the first.

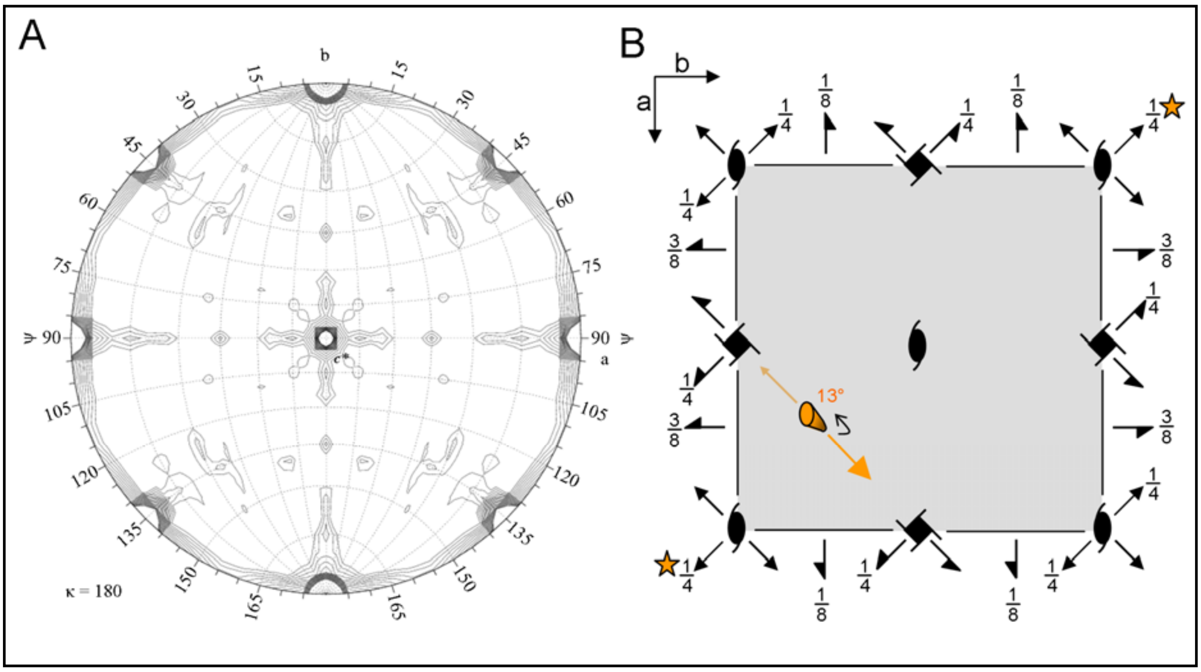
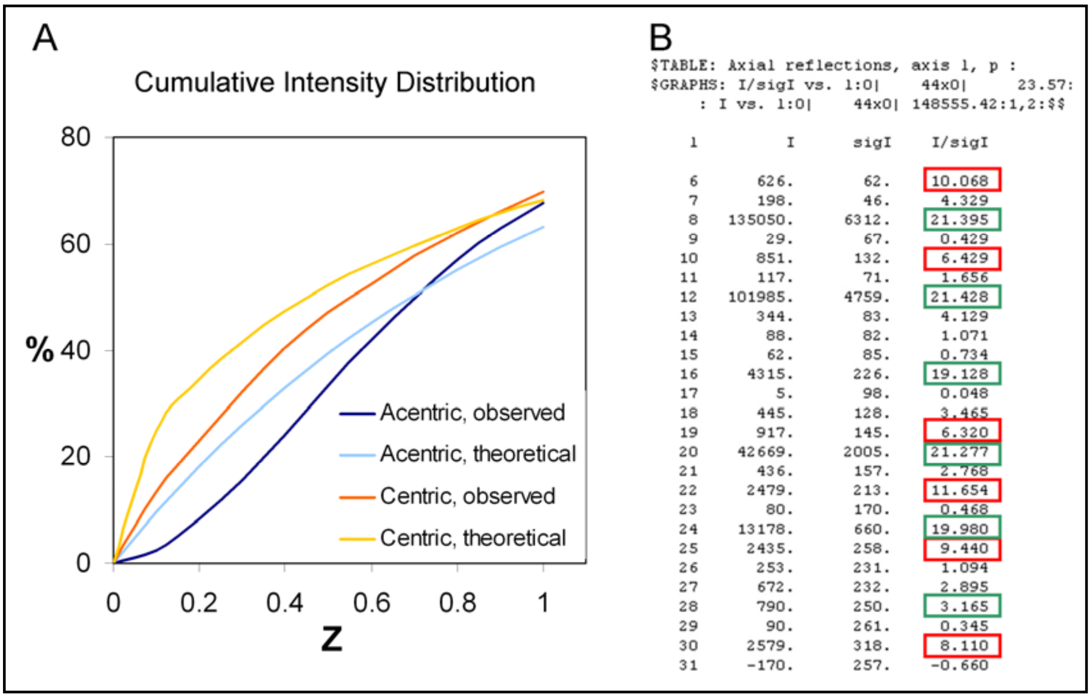


Fig. 4-11: Symmetry axes in pro-Auto A83S/A84S crystals of apparent space group $P4_32_12$. (A) Self-rotation function plot for κ = 180, showing all crystallographic two-fold axes. Peak heights are indicated by contour lines. The non-crystallographic (NCS-) axis is masked in this plot, because it runs roughly parallel to *c*. **(B)** The non-crystallographic two-fold rotational axis (orange rod) runs roughly through the point (¾, ¼, ¼). It is inclined by 13° away from the *c*-axis towards the point (½, 0, 0). This NCS-axis is perpendicular to the crystallographic two-fold rotational axis *a-b* (orange asterisk) that is parallel to the *ab*-plane and passes through (1, 0, *c/4)*. These two mutually perpendicular two-fold axes generate a third perpendicular two-fold axis (orange arrows), inclined by 13° relative to the *ab*-plane.

Twinning

During data analysis, several anomalies appeared, all of which were possible indicators of merohedral crystal twinning. The distribution of cumulative intensity of acentric reflections is sigmoidal, indicating statistically unusually few weak reflections (Fig. 4-12A). This is a typical sign for merohedral twinning, as each apparent reflection is actually the sum of two individual reflections from two related lattices, lowering the likelihood of weak intensities (Yeates, 1997). The twinning test routine of Xtriage (Zwart *et al.*, 2005) calculated the twin fraction to be 0.46, indicative of nearly perfect merohedral twinning in the true space group $P4_3$ or pseudomerohedral twinning in lower space groups. MAD phasing was successful in space groups with a 4_3-axis but failed in lower symmetry or with a different screw axis along the *c*-axis. Anomalously, however, the systematic absences along the *c*-axis are inconsistent with the pattern of a $4_1/4_3$ screw axis, contradicting the space group $P4_32_12$ (Fig. 4-12B).



```
$TABLE: Axial reflections, axis l, p :
$GRAPHS: I/sigI vs. l:0|     44x0|      23.57:
      : I vs. l:0|     44x0| 148555.42:1,2:$$

      l           I        sigI     I/sigI

      6        626.         62.     10.068
      7        198.         46.      4.329
      8     135050.       6312.     21.395
      9         29.         67.      0.429
     10        851.        132.      6.429
     11        117.         71.      1.656
     12     101985.       4759.     21.428
     13        344.         83.      4.129
     14         88.         82.      1.071
     15         62.         85.      0.734
     16       4315.        226.     19.128
     17          5.         98.      0.048
     18        445.        128.      3.465
     19        917.        145.      6.320
     20      42669.       2005.     21.277
     21        436.        157.      2.768
     22       2479.        213.     11.654
     23         80.        170.      0.468
     24      13178.        660.     19.980
     25       2435.        258.      9.440
     26        253.        231.      1.094
     27        672.        232.      2.895
     28        790.        250.      3.165
     29         90.        261.      0.345
     30       2579.        318.      8.110
     31       -170.        257.     -0.660
```

Fig. 4-12: Indications for twinning in diffraction data from pro-Auto A83S/A84S crystals. (A) The observed cumulative distribution of normalised acentric intensities (dark blue), deviates in a sigmoidal shape from the theoretical curve (light blue). **(B)** Output from the program scalepack, indicating inconsistencies in systematic absences along the $4_3/4_1$ screw axis. Reflections supposed to be extinct but measured with significant intensity are marked with red boxes. Green boxes mark reflections that fulfil the reflection condition l = 4n.

One of three monomers in the asymmetric unit – though consistent with experimentally determined selenium sites – is significantly less well defined in the electron density than the other two monomers. Its location within the asymmetric unit is such that it overlaps with its own crystallographic symmetry mate, implying that only one of the two positions can be filled in each asymmetric unit (Fig. 4-13). This inconsistency may have affected the overall quality of the diffraction data.

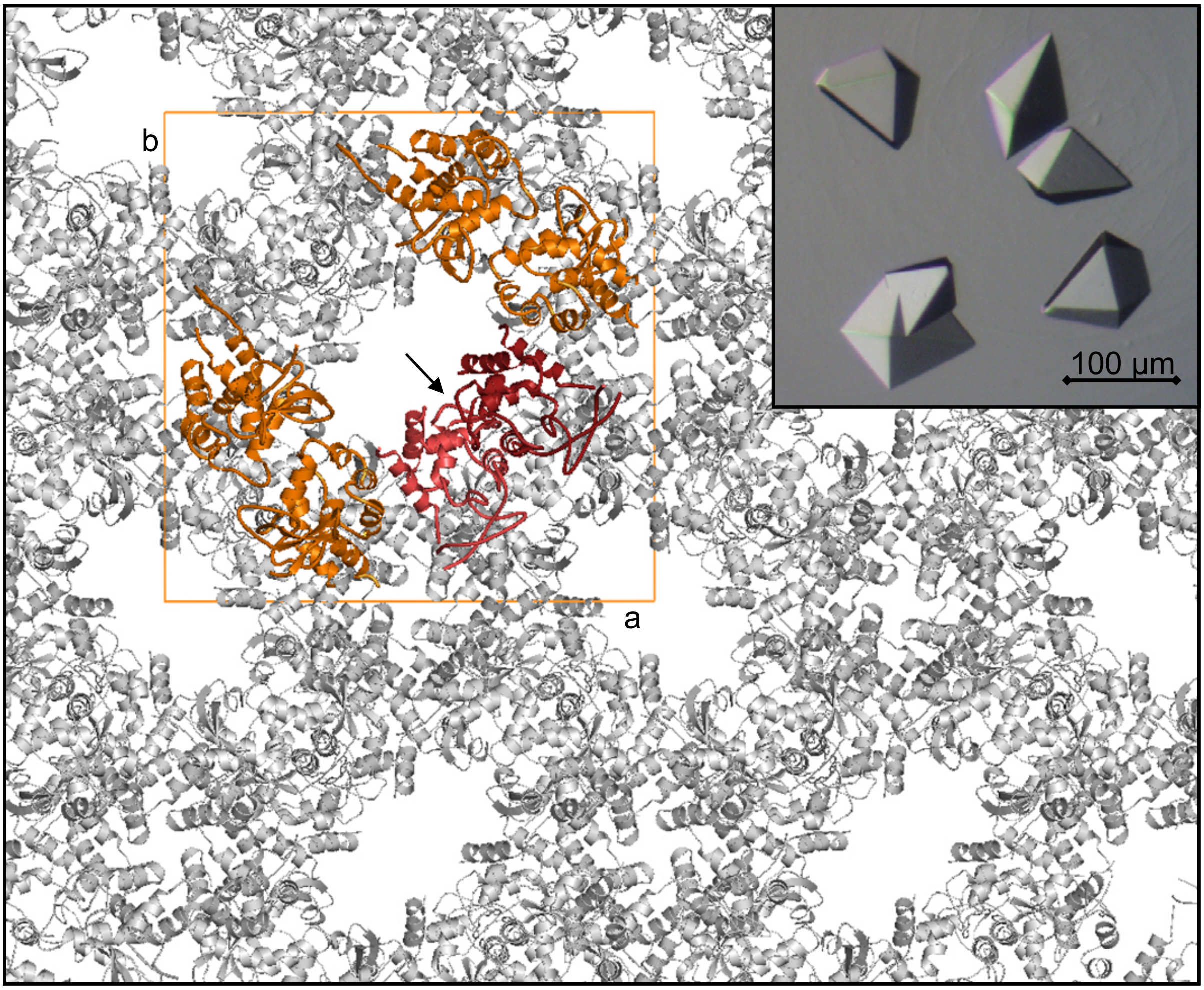


Fig. 4-13: Crystal morphology (inset) and crystal packing arrangement of pro-Auto A83S/A84S in space group $P4_32_12$, as viewed along *c*. Molecules of two asymmetric units are coloured. Orange: NCS-related molecules 1 and 2; Red: mutually exclusive third molecules. The arrow indicates the site where the third molecule clashes with its symmetry mate.

However, reindexing and scaling the data in space groups of lower symmetry and refining the structure as a case of perfect twinning in SHELXL or CNS did not significantly improve the refinement statistics, such that the R_{free}-value stalled at 35 %. Despite a sensible three-dimensional model of pro-Auto, an R_{free} of this magnitude is generally not acceptable for a 2.3 Å resolution structure, as the model can not be refined to yield expected geometric quality criteria. Since twin-correction algorithms such as DETWIN (Collaborative Computational

Project, Number 4, 1994; Taylor & Leslie, 1998) are not able to produce accurately de-twinned intensities for highly twinned crystals (Hillig & Renault, 2006), no available mathematical approach could resolve this situation. The twinned tetragonal crystal form was, therefore, abandoned and rational modifications of pro-Auto were undertaken, to obtain untwinned diffraction data.

4.1.9 Rational design of pro-Auto variants

To enable the molecular model derived from the MAD experiment to be refined to acceptable geometry and crystallographic statistics, an alternative dataset from more regular and untwinned crystals was required. The tetragonal pyramids grown in ethanol were, however, the only crystals obtained for wild-type pro-Auto and its variant A83S/A84S. All crystals displayed the same anomalies in their diffraction patterns. A set of pro-Auto variants was, therefore, designed, based on the already available structural information, to obtain an unrelated crystal packing arrangement.

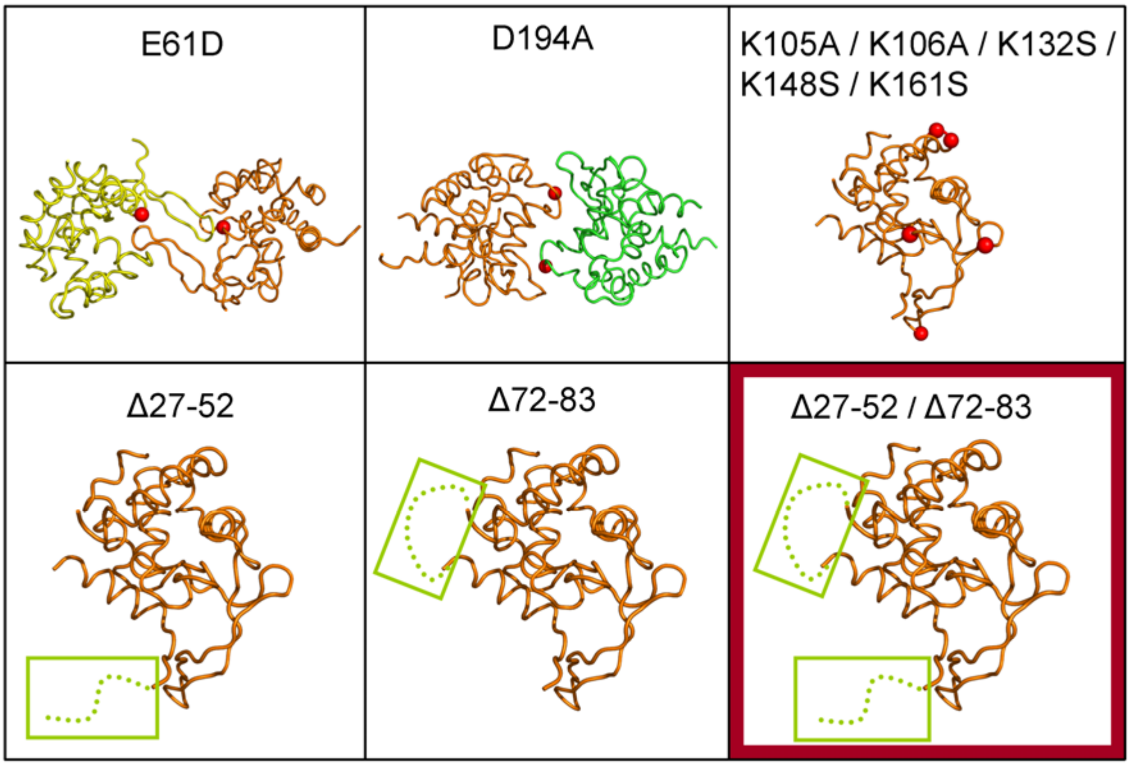


Fig. 4-14: Rationally designed pro-Auto variants to induce novel crystal packing arrangements. The MAD-derived structural model of pro-Auto was used to identify both individual amino acid substitutions and deletions of flexible regions, to disrupt the original crystal contacts and thereby potentially allow for novel crystal packing arrangements. E61D and D194A: removing crystal contacts of the twinned crystals; K105A/K106A/K132S/K148S/K161S: surface entropy reduction; Δ27-52: deletion of flexible N-terminus; Δ72-83: deletion of flexible loop (including the *activation region*); Δ27-52/Δ72-83: combination of deletions. The combination of both deletion variants (red box) resulted in a second, untwinned crystal form.

Two single amino acid substitution variants, E61D and D194A, were prepared to disrupt the main crystal contacts in the tetragonal packing. In a second strategy, five exposed lysine residues were replaced by alanine (Lys106, Lys107) or serine (Lys132, Lys148, Lys161) to decrease the surface entropy and enhance the crystallisability (Derewenda, 2004). The third approach involved shortening the disordered loop between residues 75 and 88 by deleting residues 72-82 (variant Δ72-83) and deleting the disordered N-terminal region (variant Δ27-51).

The pro-Auto variants E61D, D194A, K105A/K106A/K132S/K148S/K161S and Δ27-52 could not be crystallized. Wedge-shaped crystals obtained from variant Δ72-83, although of reasonable size (~ 50x100x50 µm) did not diffract X-rays. Only the combination of the two deletions, Δ27-51/Δ72-83, proved successful. Cigar-shaped hexagonal crystals of the combined variant grew in 1 M $(NH_4)_2SO_4$ (see inset in Fig. 4-15).

4.1.10 Crystal structure determination of the pro-Auto variant Δ27-51/Δ72-83

The variant Δ27-51/Δ72-83 was generated by removing two disordered regions in the initial, twinned-data model, allowing for an alternative crystal packing arrangement. Hexagonal crystals of this variant were untwinned and diffracted X-rays to a resolution of 2.55 Å at beamline X11 (EMBL, DESY, Hamburg). The structure was solved by molecular replacement, using the model derived from the twinned MAD data. Apart from the deleted loops, no major conformational changes are noticeable, though the geometry of the refined model is significantly better. Whereas the loop between residues 75 and 88 was disordered in the first crystal form, well-defined electron density was present for the shortened loop of the variant.

In the hexagonal crystals, neighbouring protein molecules mutually exchange their N-terminal α-helices (α0) inducing non-physiological domain swapping (see below). Domain-swapped dimers physically stabilize the crystals, allowing for wide solvent channels parallel to the *c*-axis (Fig. 4-15) and resulting in a high solvent content of 77 %. For details of data processing and refinement see Table 4-3.

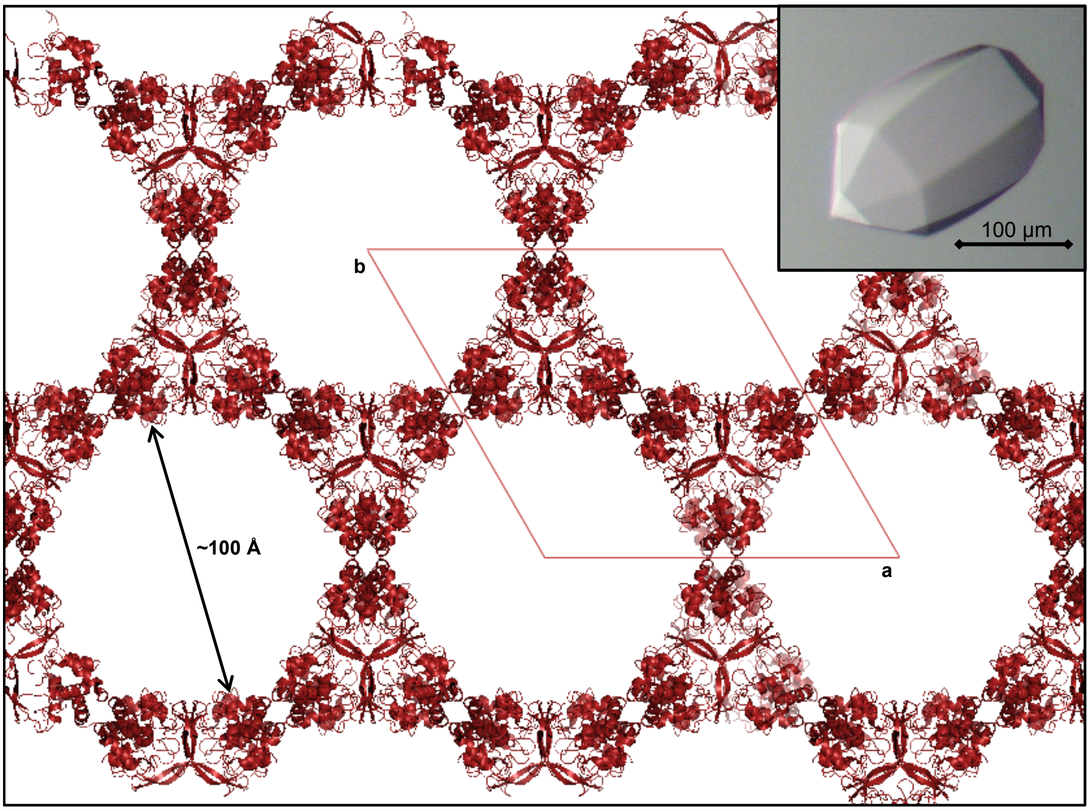


Fig. 4-15: Crystal morphology (inset) and hexagonal packing of the pro-Auto variant Δ27-51/Δ72-83, as viewed along *c*. Wide solvent channels of ~100 Å diameter account for a solvent content of 77 %.

Table 4-3: Crystallographic statistics.

Diffraction Data

Protein	**pro-Auto Δ27-52/Δ72-83**	**pro-Auto A83S / A84S selenomethionine**			
Dataset	Native	Peak	Inflection	High Remote	Low Remote
Wavelength, Å	0.81480	0.97791	0.97853	0.95369	1.03317
Space group	$P6_222$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Unit cell axes a=b, c, Å	133, 88	116, 104	115, 104	116, 105	115, 104
Resolution range, Å*	20 – 2.55	20 – 2.35	20 – 2.35	20 – 2.30	20 – 2.30
	(2.64 – 2.55)	(2.48 – 2.35)	(2.48 – 2.35)	(2.43 – 2.30)	(2.43 – 2.30)
Unique reflections*	15470 (1457)	29659 (4102)	29469 (3972)	31921 (4411)	31076 (4171)
Completeness (%)*	100 (97)	99 (96)	99 (93)	99 (95)	99 (93)
Redundancy	8	$7^{\S}$	$6^{\S}$	$8^{\S}$	16
Wilson B-factor, Å^2	58.5	46.8	49.4	43.5	48.9
I / σI*	30.4 (3.0)	21.5 (4.8)	21.7 (4.8)	18.6 (3.9)	25 (5.7)
R_{merge}, %*	8.3 (60.1)	9.2 (46.1)	8.6 (45.3)	9.4 (50)	7.4 (36.7)
Refinement statistics					
R_{work} / R_{free} (%)*	18.9 / 23.0	29.2 / 35.3			
r.m.s.d. bonds Å/angles, °	0.017/1.64	0.012/1.515			
# atoms: protein/water	1462/126	4334/53			
Ramachandran plot					
Preferred (%/#)	91.3 / 147	89 / 466			
Allowed (%/#)	8.7 / 14	9 / 47			
Outliers (%/#)	-	2 / 11			

*Values in parentheses indicate last resolution shell

§Anomalous redundancy

4.1.11 The structure of pro-Auto reveals a lysozyme-like fold

Structurally, the catalytic domain of Auto is related to lysozyme and consists of two distinct subdomains (Fig. 4-16A).

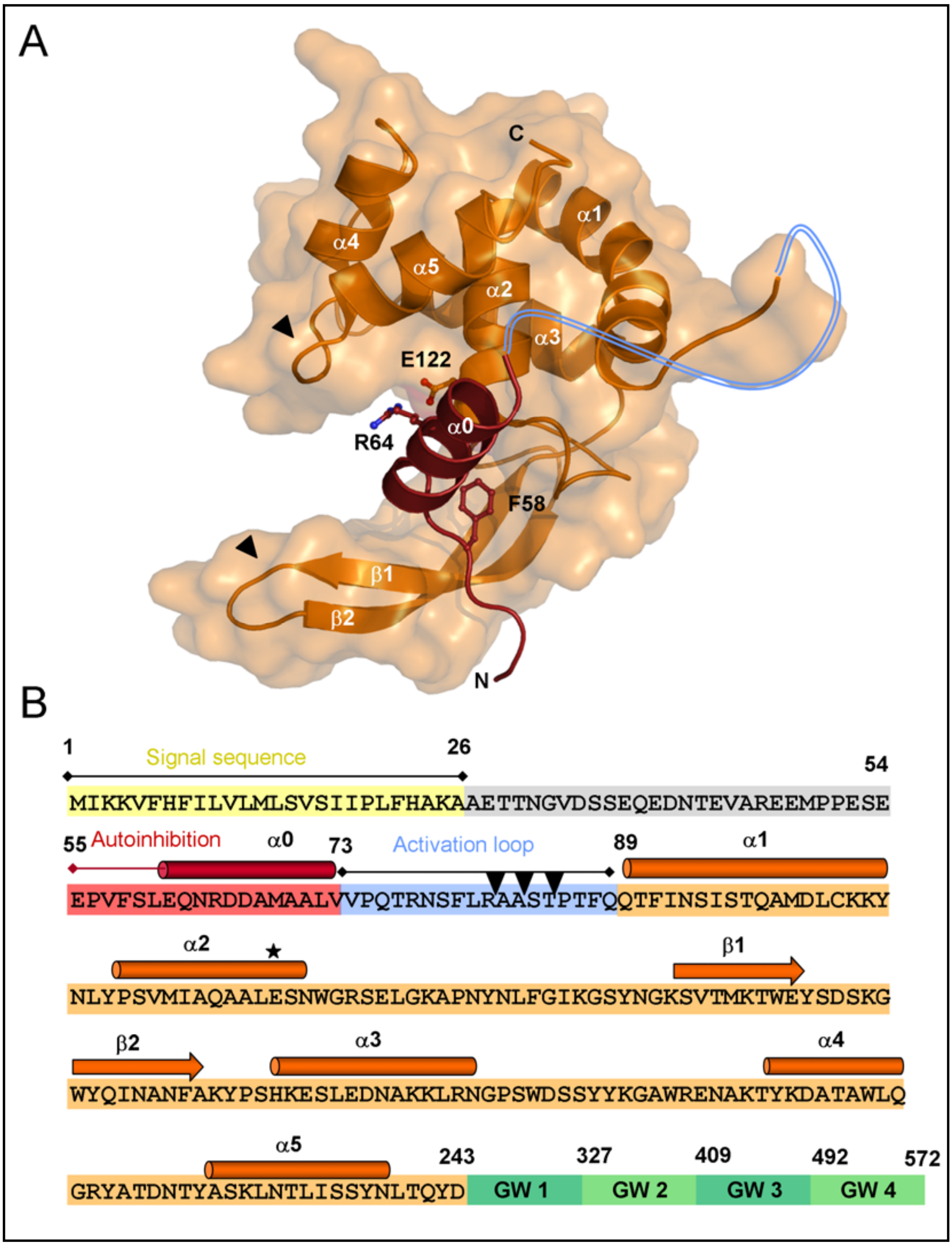


Fig. 4-16: Molecular architecture of Auto. (A) Cartoon representation of the autoinhibited catalytic domain of Auto (pro-Auto). The autoinhibitory N-terminal plug (red) fills and blocks a large part of the putative substrate binding groove. The shielded catalytic residue Glu122 and inhibitory plug residues Arg64 and Phe58 are depicted in ball-and-stick. Black triangles indicate regions with solvent-exposed aromatic residues absent from other lysozyme-like structures. **(B)** Domain structure of Auto. Secondary structure elements are marked as cylinders (α-helices) and arrows (β-sheets). Cleavage sites within the activation loop are marked with black triangles. The 'catalytic' Glu122 is marked with an asterisk.

The larger subdomain is predominantly α-helical and bears the α-helices α1 – α5. An extended β-hairpin (β1 and β2) following α2, constitutes the smaller subdomain, which, by extending away from the larger subdomain in a thumb-like manner, creates a long and deep groove, to bind the substrate.

Auto deviates from all known lysozyme-like structures by possessing an N-terminal autoinhibitory pro-sequence, which provides an additional α-helix α0 and an adjacent extended loop that fill about one half of the interdomain groove.

Despite sequence identities below 15 %, structurally related domains include the catalytic domain of Slt70 (c-Slt70) from *E. coli* (r.m.s.d. of common C_α = 2.8 Å) as well as egg-white lysozymes from goose (GEWL, r.m.s.d. = 3.2 Å) and chicken (HEWL, r.m.s.d. = 3.3 Å) and family GH19 chitinases from plants and streptomycetes (Fig. 4-17, Krissinel & Henrick, 2004).

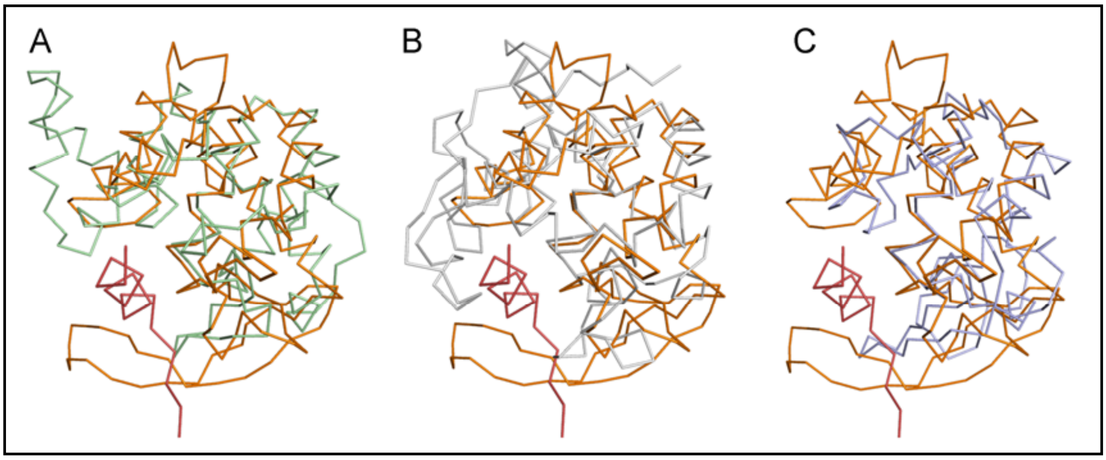


Fig. 4-17: Superposition of pro-Auto Δ27-51/Δ72-83 (orange, red) with three related structures: c-Slt70 from *E. coli* (A, PDB code 1qte), goose egg-white lysozyme (B, PDB code 153l) and hen egg-white lysozyme (C, PDB code 3lzt). The additional α-helix α0 of Auto Δ27-51/Δ72-83, absent from the related structures, is highlighted in red.

The α-helices α1-α5 of Auto are conserved in c-Slt70, GEWL and the chitinases, their lengths, however, vary substantially. HEWL is more distantly related and shares only α1, α2 and α3 with Auto.

The β-sheet that constitutes the smaller subdomain of pro-Auto, C-terminal to α-helix α2, is positionally conserved in the related enzymes. In chitinases and lysozymes this subdomain consists of three short β-strands. In Auto this is reduced to two significantly longer β-strands that create an extended β-hairpin. As a result, this subdomain protrudes prominently beyond the otherwise globular structure. A high proportion of surface-exposed residues in the β-hairpin are aromatic amino acids (Trp155, Tyr157, Trp163, Tyr164) potentially providing a suitable binding region for an oligosaccharide ligand. A further aromatic amino acid, Trp193, in the α-helical domain of Auto, may serve the same purpose.

In contrast to the evidently lysozyme-like overall fold, pro-Auto is unique in possessing α-helix α0, acting as an autoinhibitory plug. Lysozymes and Slt70 are released in an active form and do not need to be activated by proteolytic processing. In pro-Auto, the activation region is located towards the C-terminal end of a long disordered loop connecting the autoinhibitory α-helix α0 and the catalytic domain, exposing it to proteolytic cleavage. In the variant Δ27-51/Δ72-83, this loop was shortened by twelve residues, removing its structural flexibility. Coincidentally, this results in the exchange of α-helices α0 between two neighbouring molecules, forming domain-swapped dimers in the crystal (Fig. 4-18). In solution, both pro-Auto wt and the variant Δ27-51/Δ72-83 are monomeric.

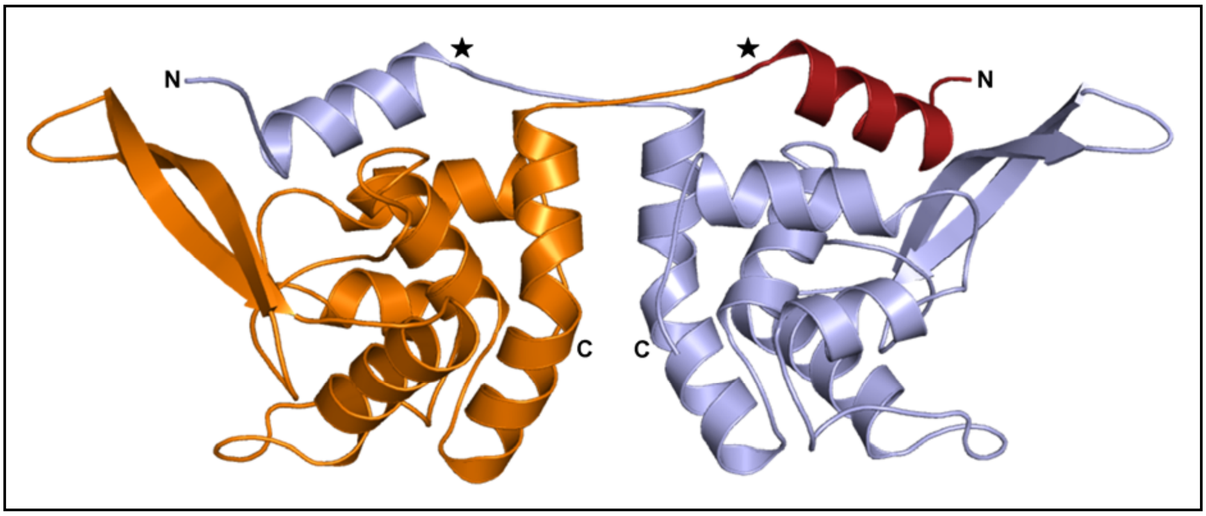


Fig. 4-18: Domain-swapped dimer in crystals of pro-Auto Δ27-51/Δ72-83. Two molecules exchange their N-terminal autoinhibitory plugs (α0). The asterisks indicate the position of an additional α-helical turn in A83S/A84S.

4.1.12 The active site of Auto contains an invariant glutamate

The active site of Auto is located within the elongated groove between the two sub-domains, supported by the fact that this groove is conserved in GEWL and c-Slt70 (Fig. 4-19). Although residues involved in saccharide-binding are not similarly conserved between Auto, GEWL and c-Slt70, exposed aromatic residues in the β-hairpin and the physical length of the binding groove indicate that multiple carbohydrate residues from a cross-linked, bulky substrate could bind. The lacking third β-strand of the smaller subdomain of Auto and the larger angle between the two subdomains causes the binding groove to be less confined than in the related structures.

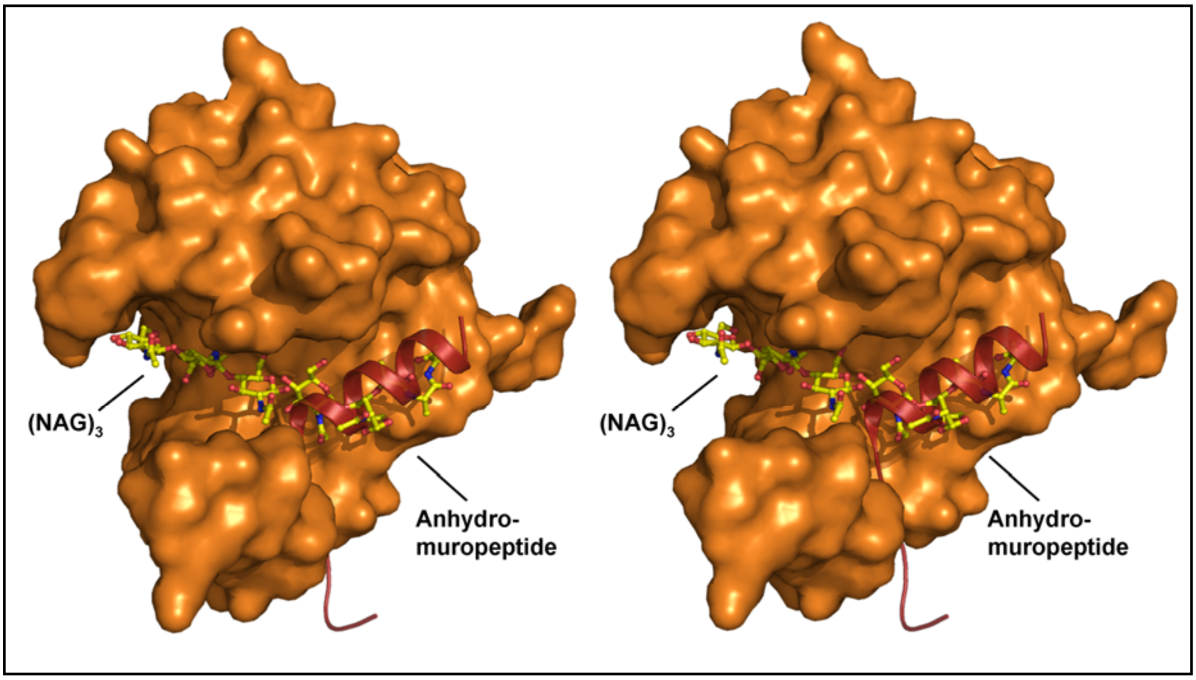


Fig. 4-19: Stereo diagram of the substrate binding groove. Surface of Auto (orange, α0 omitted) showing the possible location of substrates/inhibitors (NAG)3 and anhydromuropeptide. These were placed into the substrate binding groove by superpositioning appropriate complexes of GEWL and c-Slt70 onto pro-Auto. The N-terminal helix α0 of pro-Auto (red, transparent cartoon) occupies half of the substrate binding groove.

A conserved glutamate residue (Glu73 in GEWL, Glu478 in c-Slt70 and Glu35 in HEWL), indispensable in the catalysis of lysozymes, chitinases and soluble lytic transglycosylases, is similarly conserved in Auto in the form of Glu122. This glutamate proposedly functions as general acid by protonating the scissile glycosidic bond to initiate its cleavage (Andersen *et al.*, 1997; Grütter *et al.*, 1983; Kawamura *et al.*, 2006; Kuroki *et al.*, 1999; Malcolm *et al.*, 1989; Thunnissen *et al.*, 1994; Weaver *et al.*, 1995). Replacing Glu122 by glutamine in

Auto$_{act}$, correspondingly, abolishes enzymatic activity, confirming its equivalence to the 'catalytic' glutamates in related enzymes (Fig. 4-20).

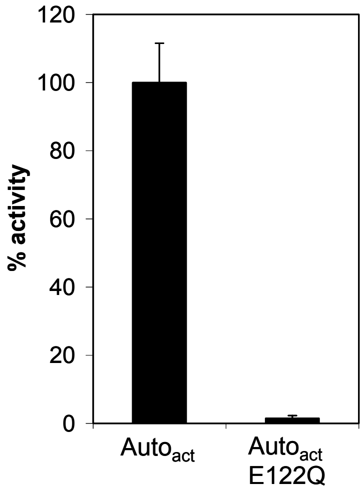


Fig. 4-20: Lytic activities of wild type Auto$_{act}$ and the variant E122Q against *M. lysodeikticus* in solution. Replacing Glu122 by glutamine abolishes activity, confirming its involvement in catalysis.

In Auto, two tyrosines and a phenylalanine surround Glu122 (compared to two tyrosines and a methionine in GEWL and c-Slt70) creating a hydrophobic environment presumably to increase its pK$_a$ and ensure protonation of the catalytic glutamate carboxyl group at higher pH (Fig. 4-21).

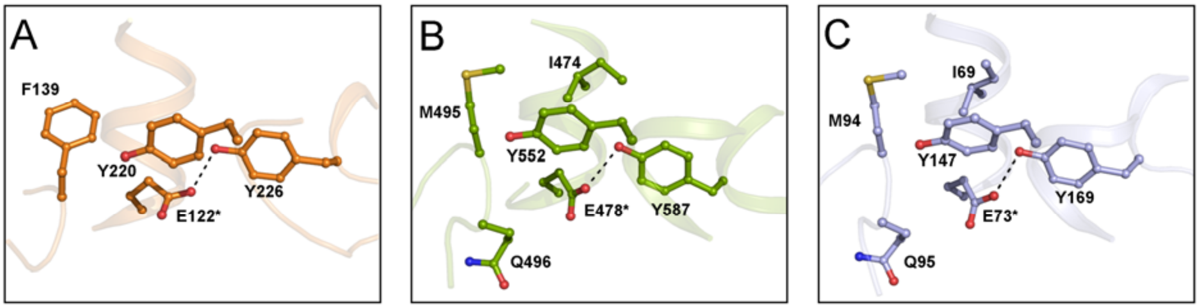


Fig. 4-21: The active sites of (A) Auto$_{act}$, (B) soluble lytic transglycosylase Slt70 from *E. coli* (PDB code 1qte) and (C) goose egg-white lysozyme (PDB code 153l). The conserved catalytic Glu, marked by an asterisk, and surrounding residues are shown in ball-and-stick.

In GEWL and c-SLT70, a spatially proximal isoleucine and a glutamine C$_\beta$ within the β-turn subdomain contribute to the hydrophobic environment. By comparison, the catalytic glutamate of Auto is more exposed, probably reducing the shift in pK$_a$ and lowering its pH-optimum. A hydrogen bond between one of the conserved tyrosines (Tyr226) and the catalytic glutamate carboxyl of Auto, again conserved in GEWL and c-Slt70, indicates that Tyr226 may additionally have a role in correctly positioning the catalytic glutamate within the active site.

4.1.13 Autoinhibition

By shortening the activation loop, the variant pro-Auto Δ27-51/Δ72-83 is constitutively autoinhibited. It is not cleaved by trypsin, chymotrypsin, thermolysin or papain, whereas subtilisin efficiently degrades the protein, as noted for the wild type protein (see Fig. 4-4). The catalytic domain of Auto is thus remarkably protease resistant, efficiently inhibited by the N-terminal pro-sequence, and is activated upon proteolytic cleavage of the activation region.

The autoinhibitory plug binds to the catalytic domain through numerous interactions (Fig. 4-22), burying a total surface area of 868 Å^2 (PISA service, European Bioinformatics Institute).

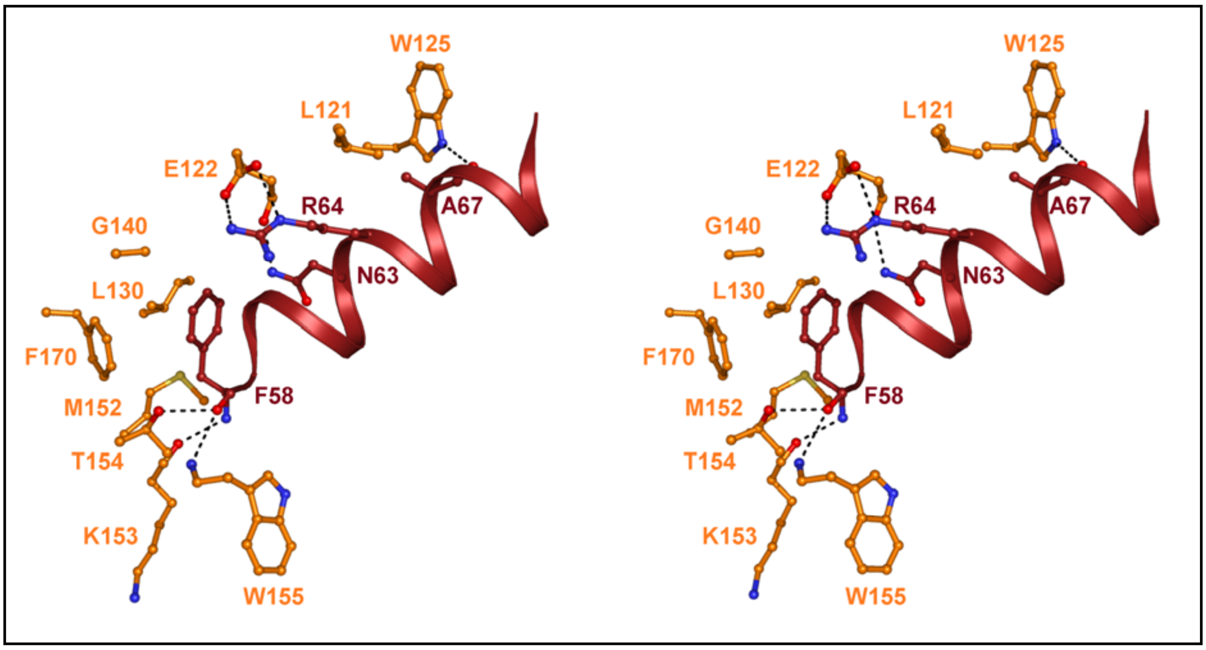


Fig. 4-22: Structural basis for autoinhibition. Stereo diagram of the interactions between the N-terminal plug and the catalytic domain. Residues mediating interactions are shown in ball-and-stick. Hydrogen bonds and salt bridges are indicated as dotted lines. Hydrophobic interactions are mediated by Phe58 and Ala67.

A central interaction occurs between Arg64 and the catalytic Glu122. Depending on the protonation status of the glutamate, this interaction could occur through a salt bridge or hydrogen bonds. Replacing Arg64 by aspartate, however, does not abolish autoinhibition.

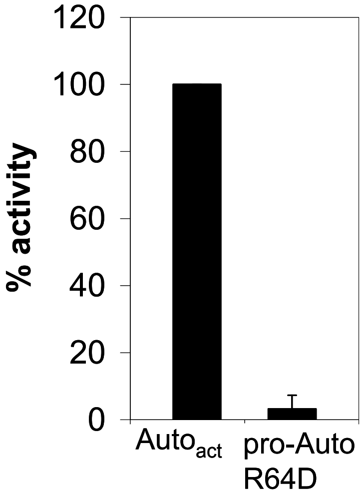


Fig. 4-23: Lytic activities of wild type Auto$_{act}$ and the variant R64D against *M. lysodeikticus* in solution. Replacing Arg64 by aspartate does not abolish autoinhibition, indicating that its interaction with the catalytic Glu122 is not indispensable for binding of the autoinhibitory plug.

This observation could indicate that the Arg64-Glu122 interaction does not significantly contribute to the binding affinity of the plug. Alternatively, the aspartate side chain may be too short to induce an electrostatic repulsion or it may allow a water- or cation-mediated interaction to Glu122. In view of the numerous additional interactions between plug and catalytic domain, the function of Arg64 may primarily be to shield Glu122 rather than dramatically affecting binding affinity. Hydrogen bonds link the backbone of Phe58 to Lys153, Trp155 and Thr154, as well as Asn63 to the carbonyl of Glu122, and the carbonyl oxygen of Ala67 to Trp125. Two hydrophobic interactions are involve Phe58 inserting into a hydrophobic pocket within the substrate binding groove created by Leu130, Gly140, Met152, and Phe170, and Ala67 interacting with Leu121 (Fig. 4-22, Fig. 4-23).

The length of the inhibitory α-helix represents the only major structural difference between the pro-Auto variants A83S/A84S and Δ27-51/Δ72-83 involving the main-chain. In the domain-swapped variant with the shortened loop, the helix consists of 2.5 turns, compared to 3.5 turns in A83S/A84S (Fig. 4-18). It is not clear whether this partial unwinding is due to a tensile force caused by shortening the connection between α-helix and the catalytic domain or due to the observed swapping of these helices between neighbouring Auto monomers.

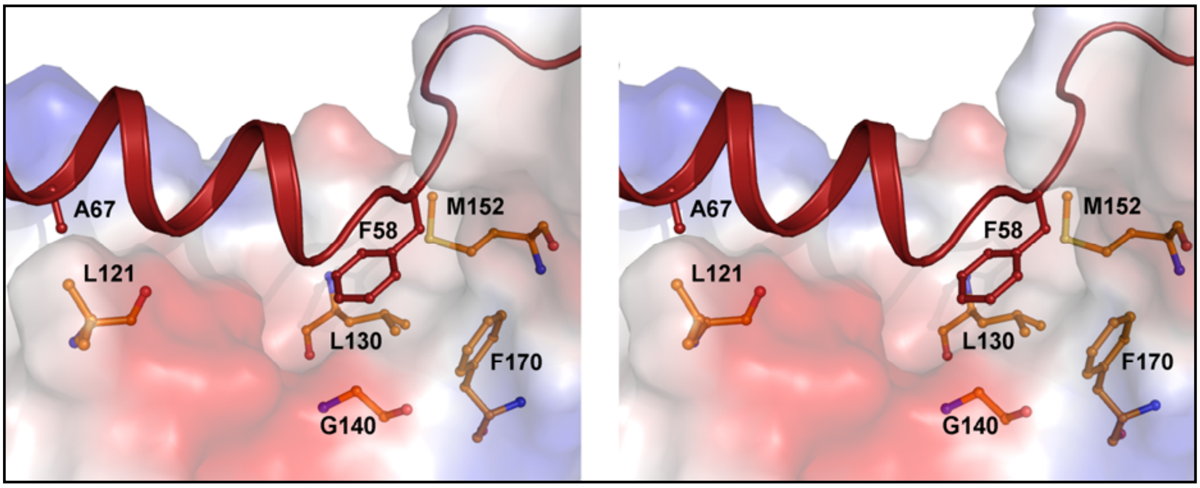


Fig. 4-24: Stereo diagram depicting the hydrophobic interactions between the inhibitory plug and the substrate binding groove. Phe58 inserts into a hydrophobic pocket created by Leu130, Gly140, Met152, and Phe170. Ala67 closely faces Leu121. The electrostatic surface potential of the substrate binding region is shown: red: negatively charged, blue: positively charged, white: uncharged.

4.2 InlJ

This study involved the structural characterization of the cysteine-rich internalin domain of InlJ, henceforth referred to as InlJ'. InlJ' was cloned and recombinantly produced in *E. coli*. It was then purified chromatographically and analysed by X-ray crystallography. By comparing its primary and tertiary structures to those of related proteins, allowed novel and unique features of this unusual type of cysteine-rich LRR protein to be identified.

4.2.1 Bacterial cysteine-containing (bCC) LRR-type proteins

To explore the distribution of the bacterial cysteine-containing (bCC) LRR-type, sequence data banks were searched for the motif x-L-x-x-L-x-C using BLAST. The resulting sequences were narrowed down to those containing at least 2 repeats matching the pattern. Table 4-4 lists all currently known genes of the bCC LRR-type.

InlJ is found in all sequenced *L. monocytogenes* strains: EGD-e (Glaser *et al*., 2001), F6854 (Nelson *et al*., 2004), 10403S (Birren *et al*., 2006, all serovar 1/2a), H7858 (Nelson *et al*., 2004), and F2365 (Nelson *et al*., 2004, both serovar 4b).

Two *L. monocytogenes* strains (EGD-e and 10403S) possess a further bCC-LRR protein (lmo0331, LMRG00023) of unknown function, which – in contrast to InlJ – has a homolog in *L. innocua* (Lin0354). The serovar 1/2a strain F6854 and the serovar 4b strain, furthermore, respectively encode two and one bCC-LRR proteins in addition to InlJ. The bCC-repeat is not restricted to the genus *Listeria* but is found in several other organisms, including bacterial pathogens and commensals as well as viruses.

Interestingly, an InlJ-homolog from the Gram-negative anaerobe *Porphyromonas gingivalis* has recently been identified and was shown to influence bacterial accumulation upon biofilm formation (Capestany *et al*., 2006).

The majority of bCC-LRR proteins conform to the 21-residues repeat length and appear mostly as a building block in several paralogous proteins (Sabet *et al*., 2005).

Table 4-4: Bacterial cysteine-containing (bCC) -type LRR proteins.

Name	**Accession Number(s)**	**Organism**	**# LRR**	**# bCC-LRR**	**Length**
InlJ	(*)	*L. monocytogenes* (*)	15	13	21
lmo0331, LMRG00023	NP_463861, EBA23962	*L. monocytogenes* (EGD-e, 10403S)	11	10	21
lmoh7858_0367, lmof2365_0347	ZP_00229263, YP_012956	*L. monocytogenes* (H7858, F2365)	12	11	21
lmof6854_0363	ZP_00234780	*L. monocytogenes* (F6854)	3	2	21
lmof6854_0362	ZP_00234779	*L. monocytogenes* (F6854)	8	8	21
lin0354	NP_469699	*L. innocua* (Clip11262)	11	10	21
RCC1	ZP_01244003	*Flavobacterium johnsoniae*	19	16	21
FjohDRAFT_4748	ZP_01244004	*Flavobacterium johnsoniae*	20	16	21
FjohDRAFT_1094	ZP_01247123	*Flavobacterium johnsoniae*	10	8	21
Chu0515	YP_677143	*Cytophaga hutchinsonii*	12	12	21
Chu2920	YP_679504	*Cytophaga hutchinsonii*	10	5	20-23
Chu2055	YP_678660	*Cytophaga hutchinsonii*	12	5	22
Chu0459	YP_677088	*Cytophaga hutchinsonii*	12	10	20
Chu1860	YP_678468	*Cytophaga hutchinsonii*	8	6	21-22
PG0350	NP_904662	*Porphyromonas gingivalis*	11	8	21
PG1374	NP_905539	*Porphyromonas gingivalis*	9	9	21
EF2250	NP_815909	*Enterococcus faecalis*	*5*	5	21
TDE0206	NP_970822	*Treponema denticola*	20	16	21
TDE0593	NP_971207	*Treponema denticola*	23	22	21
TDE2003	NP_972606	*Treponema denticola*	13	12	21
TDE2231	NP_972831	*Treponema denticola*	8	8	21
TDE1110	NP_971718	*Treponema denticola*	9	8	21-27
Pnap3264	YP_983483	*Polaromonas naphtalenivorans*	13	13	21
CdifQ_03001782	ZP_01634887	*Clostridium difficile*	6	3	21
MldDRAFT_4836	ZP_01289150	*Delta proteobacterium MLMS-1*	10	10	21
BT0284	NP_809197	*Bacteroides thetaiotamicron*	12	9	21-24
Pcryo0038	YP_580203	*Psychrobacter cryohalolentis*	10	3	23
MSV251	NP_048322	*Melanoplus sanguinipes entomopoxvirus*	9	9	21
AMVITR01	NP_065061	*Amsacta moorei entomopoxvirus*	18	17	22

(*): all sequenced strains: EGD-e (lmo2821; NP_466343), 10403S (EBA22480), H7858 (ZP_00231051), F2365 (YP_015399), F6854 (ZP_00233236)

4.2.2 Protein production and purification

InlJ', comprising the cap-, LRR- and interrepeat region (residues 34-508) was produced in *E. coli* with an N-terminally fused glutathione S-transferase affinity tag. The fusion protein was purified from the cytoplasmic fraction, cleaved with TEV-protease to remove the tag and concentrated to 7 mg/ml for crystallization. Typical yields amounted to ~2 mg pure InlJ' per L of culture. Table 4-5 summarizes important physico-chemical parameters of the InlJ'-construct with and without the GST tag.

Table 4-5: Physico-chemical parameters of GST-InlJ' and InlJ', as calculated with VectorNTI (Invitrogen).

Parameter	GST-InlJ'	InlJ'
Length (aa)	728	479
Molecular weight (kDa)	81.8	52.9
Molar extinction coefficient	73990	25860
Isoelectric point (pI)	4.9	4.4

4.2.3 The majority of cysteine residues in InlJ' is buried

The unusually high number of 14 cysteine residues in InlJ provoked the question of their accessibility from the solvent and their probable oxidation state. To address this issue, recombinant InlJ' was incubated with DTNB (‘Ellman’s reagent’), and the absorption at 412 nm of released TNB upon reaction of DTNB with free sulphhydryl groups was measured and compared to dilutions of L-cysteine under native (buffer) and denaturing (guanidinium hydrochloride) conditions (Fig. 4-25).

As expected, most cysteines are only observed when InlJ' has been denatured, confirming that the cysteines are buried within the hydrophobic core of the protein and are inaccessible to oxidation or modification. Quantifying the number of accessible cysteines indicates that about one cysteine reacts in native InlJ' while fifteen react in denatured InlJ'. As InlJ' only contains fourteen cysteines, the signal is clearly slightly inflated presumably due to limitations in the photometrical method. An error of only 0.1 µM in the protein concentration could account for the observed difference. Scaling the data would imply that overall one of 14 cysteines per InlJ' is partly accessible in the native state.

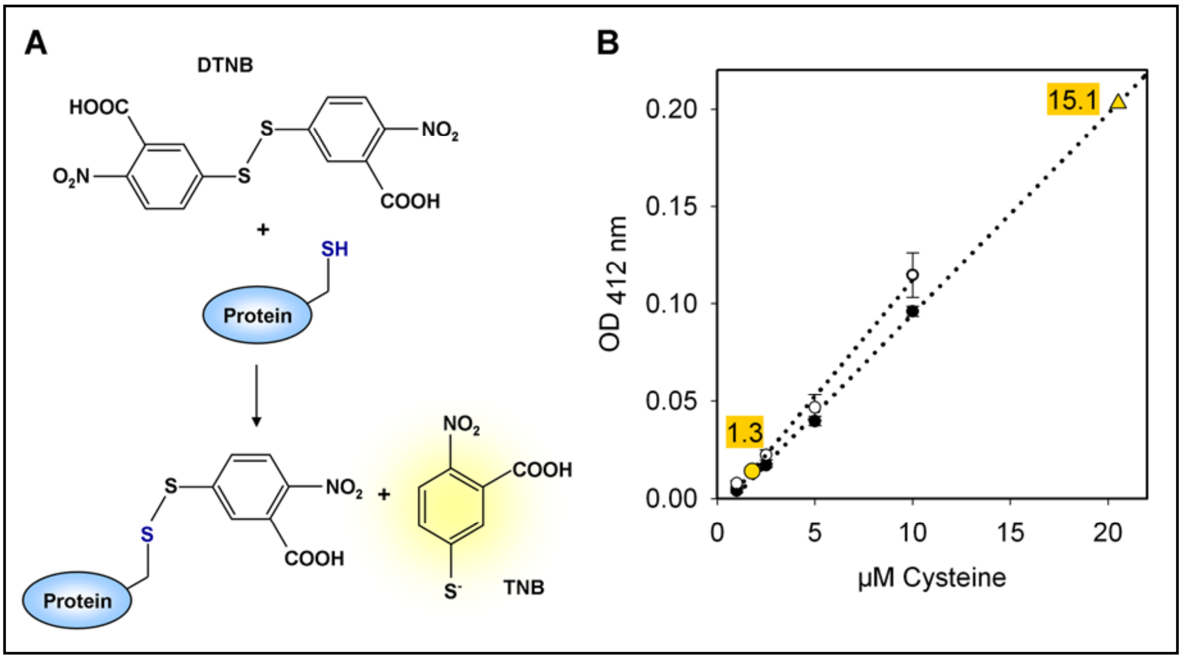


Fig. 4-25: Detection of accessible cysteines in InlJ' with DTNB. (A) Upon reaction of DTNB with accessible thiol groups of a protein, TNB is released and its absorption can be detected at 412 nm. **(B)** Accessible cysteines in InlJ'. In the native state, one cysteine per molecule InlJ' is detected. Upon denaturation with guanidinium hydrochloride, the signal increases to a calculated fifteen cysteines per molecule InlJ'. Hollow circles: calibration with L-cysteine in buffer, filled circles: calibration with L-cysteine in guanidinium hydrochloride. Dotted lines indicate regression lines of calibration measurements. Yellow dot: native signal from InlJ' on regression line (buffer), yellow triangle: signal from denatured InlJ' on regression line (guanidinium hydrochloride).

4.2.4 Crystallization, data collection and structure solution

Initial InlJ' crystals grew in a 600 nL sitting drop screening format at 20°C, using 2 M $(NH_4)_2SO_4$, 0.1 M Na acetate pH 4.6. Subsequent optimization of the crystallization conditions led to the growth of large cubic crystals at 25°C in 24-well hanging drop vapour diffusion plates containing 2 M $(NH_4)_2SO_4$, 0.1 M Na acetate pH 4.4 – 4.6 (Fig. 4-26).

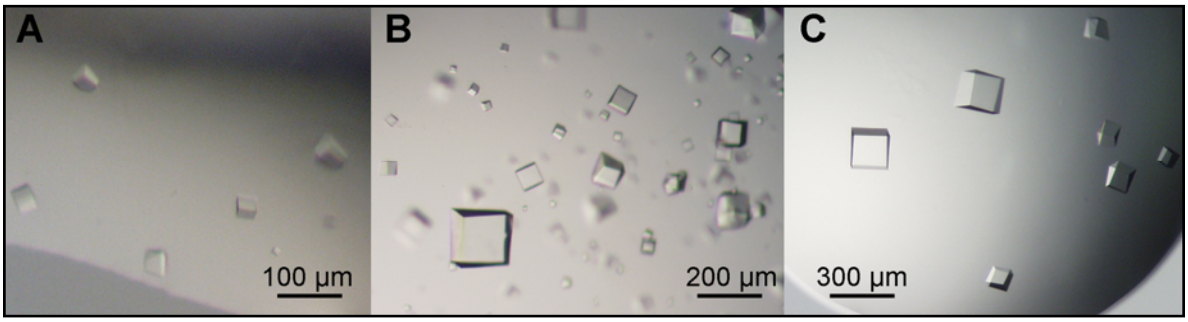


Fig. 4-26: Initial (A, 20°C) and optimized (B, C, 25°C) cubic crystals of InlJ', grown in 2 M $(NH_4)_2SO_4$, 0.1 M Na acetate pH 4.4 – 4.6.

Diffraction data were collected at beamline X11 (EMBL, DESY, Hamburg) at 100 K. Crystals belong to the cubic space group I23 with unit cell edge length 159 Å. Data were indexed, integrated and scaled using the HKL2000 suite.

A search model encompassing LRRs 1-13 of InlA (PDB code 1O6V) was generated by CHAINSAW using the sequence of InlJ'. This allowed the LRR region to be located by maximum-likelihood based molecular replacement (PHASER). The interrepeat region of InlB was then positioned in a second round. The partial models were subjected to rigid body refinement, simulated annealing combined with individual B-factor refinement and energy minimization (CNS), followed by several cycles of manual rebuilding (COOT) and restrained refinement including TLS protocols (REFMAC5). The cubic crystal packing arrangement is depicted in Fig. 4-27. Data and refinement statistics and geometric parameters of the final model are listed in Table 4-6.

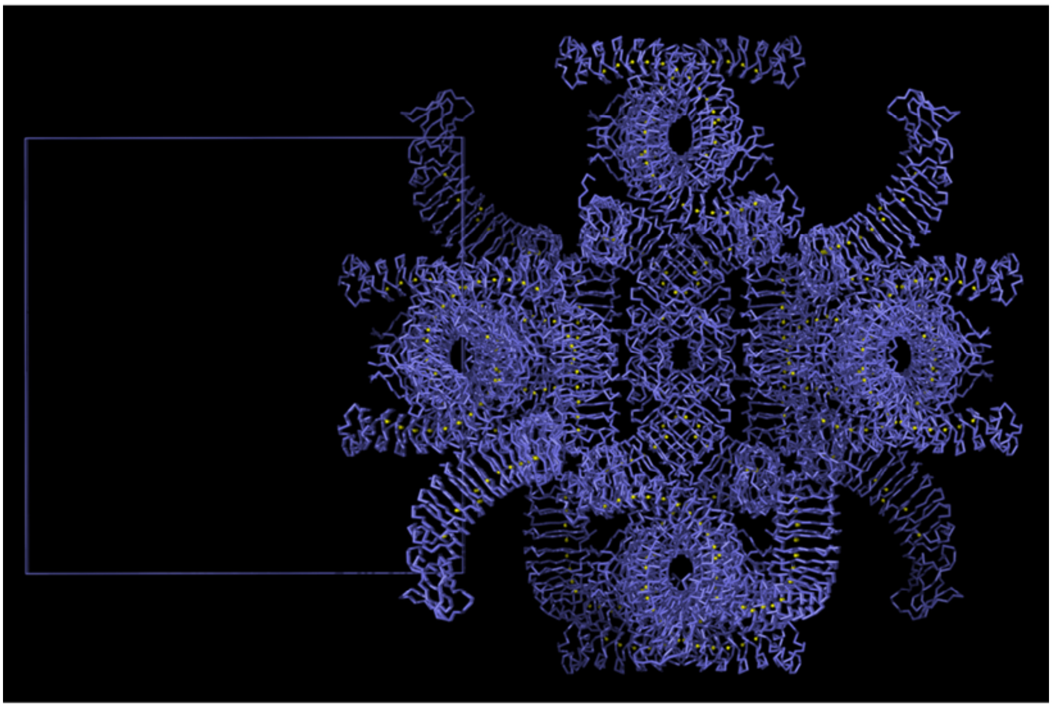


Fig. 4-27: Cubic crystal packing arrangement of InlJ'. InlJ'-molecules are depicted as blue C_α-ribbons, sulphur atoms as yellow spheres. The blue box encloses a unit cell of the crystal.

Table 4-6: Data collection and refinement statistics for InlJ'.

Diffraction Data	
Space group	I23
Unit cell dimensions	
a=b=c, Å	159.2
Wavelength, Å	0.817
Mosaicity, °	0.7
Resolution range, Å*	20 - 2.5 (2.59-2.50)
Completeness, %*	99.8 (99.7)
I/σI*	18 (2.7)
Redundancy*	12 (8)
R-merge, %*	12.1 (78)
Wilson B-factor, Å^2	49.8

Refinement statistics	
Resolution range*	20 - 2.5 (2.59-2.50)
R_{work}/ $R_{free,}$ %	20.1 / 25.1
Ramachandran plot	
Preferred, %/#	77.4 / 326
Allowed, %/#	21.6 / 91
Generously allowed, %/#	0.7 / 3
Disallowed, %/#	0.2 / 1
r.m.s.d. bonds, Å/angles, °	0.014/1.5
# atoms protein/water	3519/193

*Values in parentheses refer to the shell of highest resolution.

4.2.5 The crystal structure of InlJ'

Apart from the N-terminal signal peptide of 33 amino acids (not cloned), another 19 residues are not visible in the electron density, making Thr53 the first residue in the structural model. This truncation occurs "spontaneously" in protein samples stored for extended periods.

Overall the structure of InlJ' is similar to that of other internalin family members (Fig. 4-28A). N-terminally, the hydrophobic core is shielded from the solvent by the α-helical cap domain, C-terminally it extends into the adjacent β-sandwich of the Ig-like domain, from the tip of which the polypeptide continues to the C-terminal repeats. The C-terminus of the entire internalin domain is located at the distal tip of the Ig-like domain, ensuring the physical separation of the LRR recognition domain from the C-terminal MucBP repeats (Schubert *et al.*, 2001).

Difficulties in solving the structure of InlJ' using the entire internalin domain of InlA as a model are clear in retrospect, as InlJ and InlA differ substantially in their curvature, twist and lateral bending, particularly with respect to the last four repeats (Fig. 4-28B).

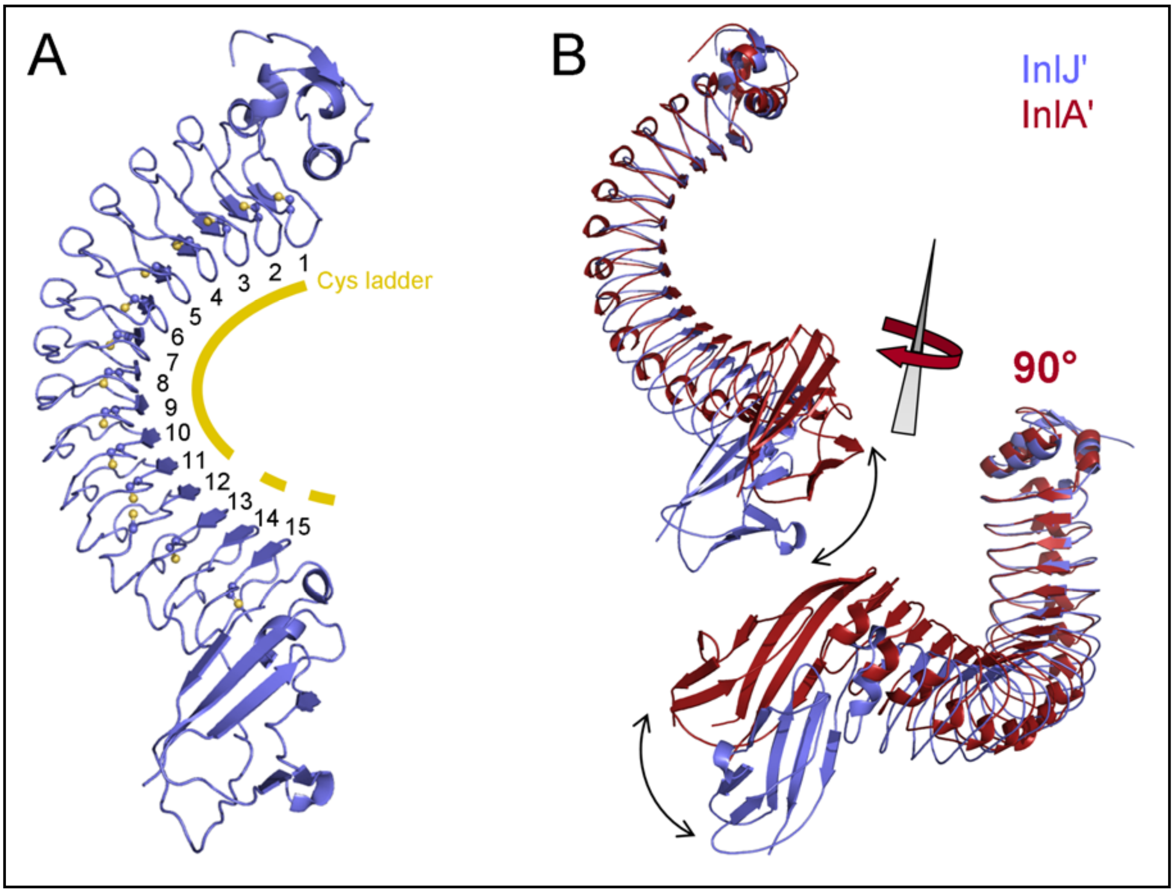


Fig. 4-28: Cartoon representation of InlJ'. (A) LRRs are numbered sequentially and cysteines are included in ball-and-stick mode. The extent of the cysteine ladder is marked by a yellow curved bar. **(B)** Superposition of InlJ' (blue) and InlA' (red) based on the C_α-atom positions in LRR1-3. Arrows indicate differences in their LRR geometries. Upper left panel: InlJ' is significantly less curved than InlA', especially in LRR12-15. Lower right panel: the same superposition but rotated by 90° around an inclined axis to demonstrate the deviation in curvature and lateral bending between InlJ' and InlA'.

The canonical LRR of InlJ comprises 21 residues, which is one residue less than the 22-residue repeat characteristic for all other internalins. In fact, the deletion of this one residue is invariably located on the outer, convex side of the solenoid. Correspondingly, in InlJ an aspartate adjacent to the deleted residue replaces a conserved proline (position 15) in other internalin family proteins (see also section 1.4.1, Fig. 1-10). This modification results in the loss of the 3_{10}-helix at the convex surface of other internalins and its substitution by a short loop conformation in InlJ'. Linear strings of deeply embedded and regularly positioned water molecules replace direct inter-repeat interactions by a tight network of hydrogen bonds along the convex side of the solenoid.

Of the fifteen LRRs of InlJ, four (LRR1, 6, 12 and 15) deviate from the 21-residue repeat length by the addition of a 22nd residue (Fig. 4-29). Only in two of these 22-residue repeats (LRR1 and 15) does this insertion at position 15 restore a classical internalin repeat re-establishing the 3_{10}-helix.

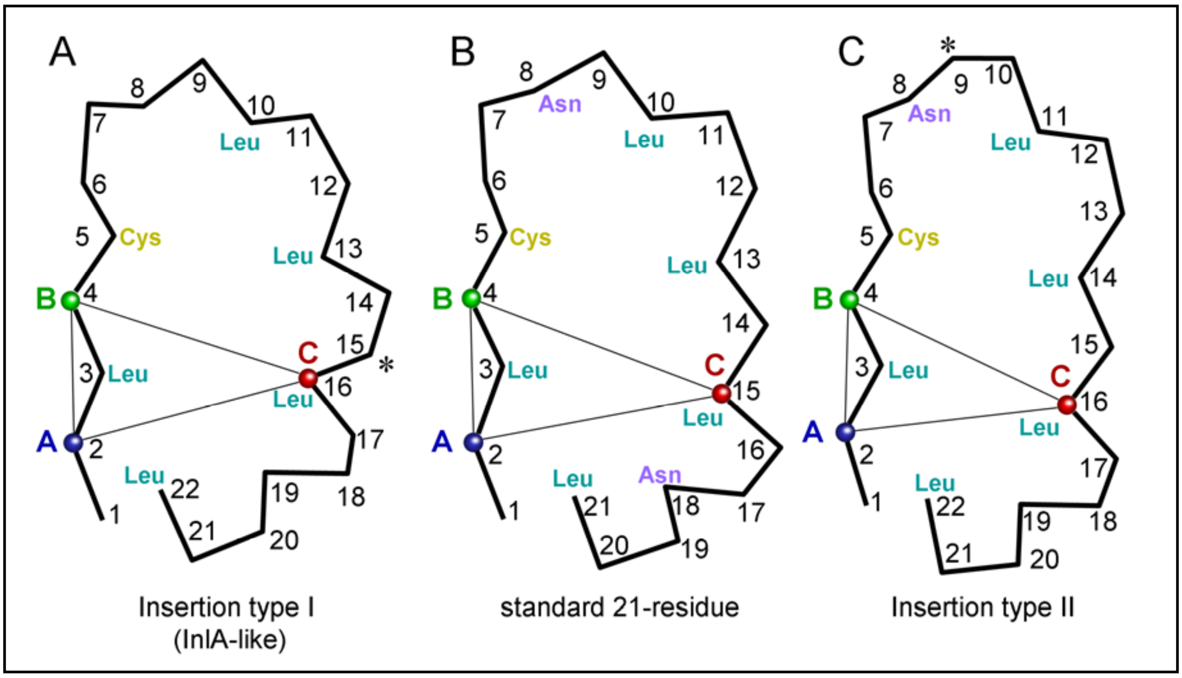


Fig. 4-29: The three LRR-types of InlJ. Residue positions within each LRR are numbered sequentially. Apart from a standard 21-residue LRR **(B)**, InlJ bears two distinct 22-residue LRRs **(A)** and **(C)**. Inserted residues relative to the 21-residue repeat are marked by asterisks. Three C_α-atoms A, B and C, used to define each repeat plane, are marked as blue, green and red spheres respectively. Core-facing residues are indicated by three-letter code of residues most common at each position.

In the remaining two 22-residue-repeats (LRR6 and 12), the additional residue is inserted at position 9 in the loop linking the concave to the convex side of the protein. This second insertion affects the overall twist of the molecule, as discussed later.

4.2.6 Cysteine and asparagine ladders in InlJ'

In addition to the Asn-ladder, involving asparagine in position 8 of many LRR proteins, InlJ possesses a previously undescribed cysteine ladder, consisting of a similarly regular arrangement of cysteines. The cysteine replaces a hydrophobic residue in position 5 of the canonical 21-residue repeat. Correspondingly, the cysteines are all located in the hydrophobic core of the protein, with each thiol group pointing towards the cysteine-C_β of the following repeat (Fig. 4-30). The only exception is Cys336 in LRR12, which is not involved in the ladder, but is shifted by three positions (towards the C-terminus), and points ‘upwards’, towards the preceding repeat, thereby defining the end of the asparagine ladder. Thus, although InlJ is an extracellular protein and is exposed to an oxidizing, aerobic environment, (especially during secretion), the cysteines within the ladder remain in their reduced form and do not form disulphide bridges.

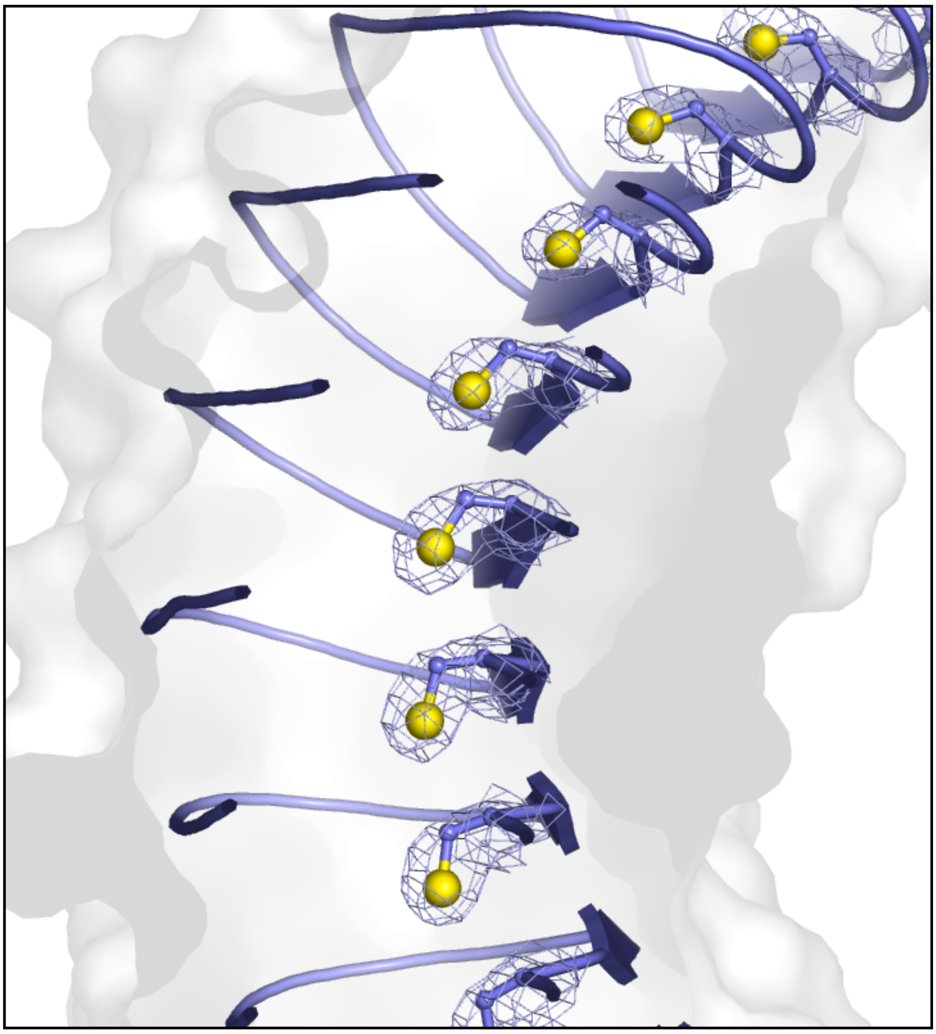


Fig. 4-30: Detailed view of the cysteine ladder in InlJ'. Cysteine residues are shown in ball-and-stick. The blue mesh depicts the electron density (2Fo-Fc at 1σ) of the cysteine residues.

4.2.7 Surface features of the InlJ'-monomer

The structure of InlJ' reveals several solvent-exposed aromatic and hydrophobic side chains at the concave surface, supposing that this region might be responsible for recognition of a specific receptor. In particular, InlJ presents at least one solvent-exposed aromatic and/or hydrophobic residue at the concave side in nine out of its fifteen leucine-rich repeat windings. Residues include Tyr140 in LRR3 (or Tyr140^3), Tyr161^4, Tyr182^5, Phe230^7, Phe267^9, Tyr288^{10}, Ile313^{11}, Tyr330^{12}, Tyr351^{13}, Tyr373^{14} and Tyr375^{14}.

Potential protein-protein binding sites in InlJ' were predicted *in silico* using the ProMate software (Neuvirth *et al.*, 2004). This analysis concurs with our interpretation that the concave surface of the LRR region is the most probable protein-protein interaction site (Fig. 4-31).

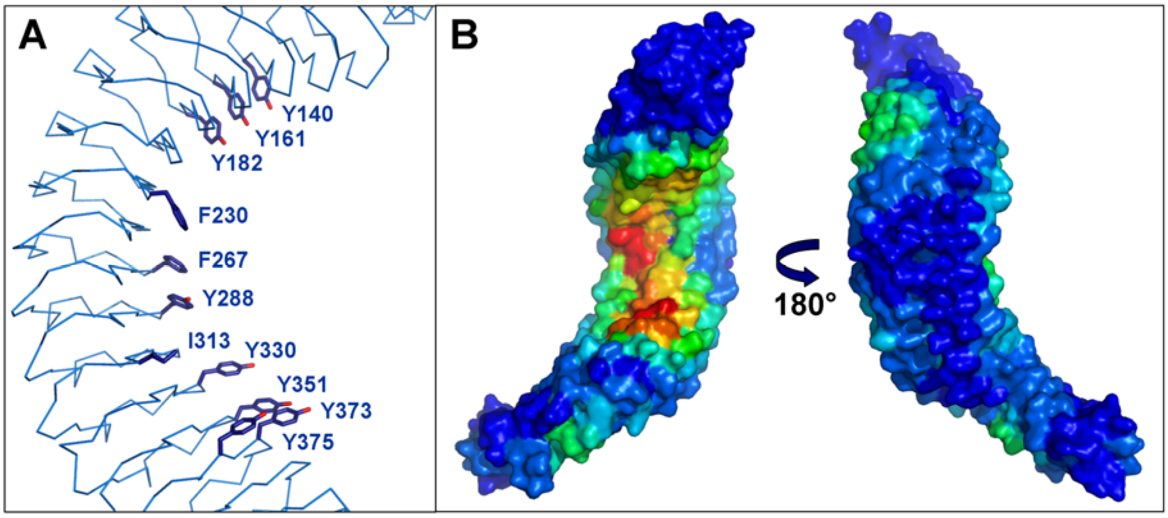


Fig. 4-31: Probable receptor-binding surface of InlJ. (A) Solvent-exposed residues enriched in protein-protein interfaces compared to other surfaces are shown in ball-and-stick mode and labelled. **(B)** Surface representation of the ProMate binding site probability score. Red: high interface probability, blue: low interface probability.

The accuracy of this prediction was cross-checked by analysing InlA and InlB, for which the interactions with human E-cadherin (hEC1) (Schubert *et al.*, 2002) and human Met-receptor (Machner *et al.*, 2003; Niemann *et al.*, 2007) are known. ProMate (Fig. 4-32) is clearly able to differentiate the interacting (concave) surface from the non-interacting (convex) surface, and is able to identify regions involved in recognition. Only the number of residues actually involved in protein-protein recognition appears to be slightly overestimated.

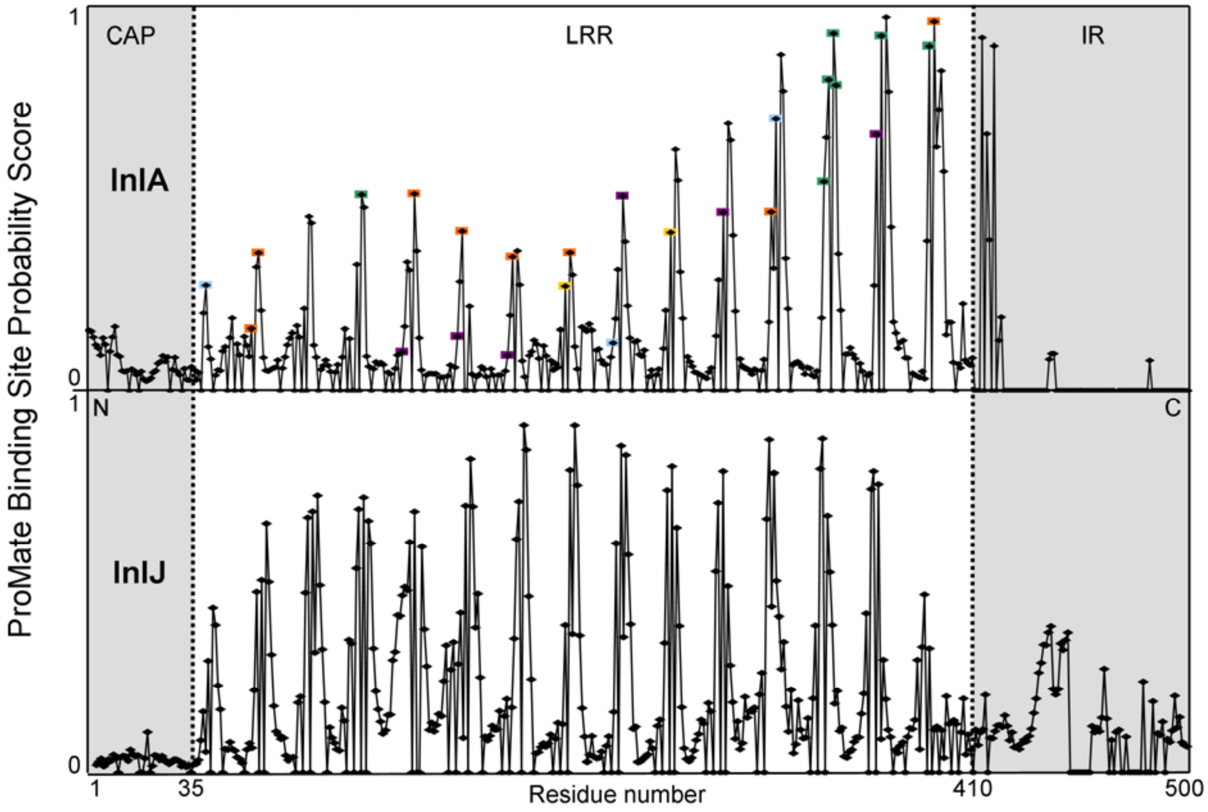


Fig. 4-32: Probability of single residues in InlJ and InlA to contribute to a protein-protein binding site, as calculated by the ProMate software. Probability scores are plotted on an arbitrary scale along the primary structure of the functional domains of InlA and InlJ. Individual residues involved in InlA / hEC1-interaction are highlighted by coloured boxes. Light blue: salt bridge, orange: hydrogen bond to bridging water molecule, green: hydrophobic interaction, purple: hydrogen bond, yellow: unfavourable interaction**.**

4.3 InlG

The gene of mature InlG, comprising the internalin domain and the variable region (residues 36-460, lacking the N-terminal cell sorting signal and ending at the C-terminal LPXTG motif) had previously been cloned into the vector *pGEX-6p-1* by Julia Ehinger (Ehinger, 2002). After unsuccessful crystallization trials, she generated a truncated construct by inserting a stop codon after residue 277, resulting in a construct of the internalin domain alone (from now on referred to as InlG'). Previous work on other internalin proteins had proven this domain to be a stable and crystallisable unit.

In this work, InlG' was produced, purified, crystallized and its crystal structure was solved at 2.05 Å resolution.

4.3.1 Protein production and purification

The protein was expressed with an N-terminal GST-tag. Its removal after the initial affinity chromatography by PreScission® protease leads to five additional residues (Gly-Pro-Leu-Gly-Ser) to be retained at the N-terminus of InlG'. Table 4-7 lists main physico-chemical parameters of the GST-fusion and the cleaved protein.

Table 4-7: Physico-chemical parameters of GST-InlG' and free InlG' as calculated with VectorNTI (Invitrogen).

Parameter	GST-InlG'	InlG'
Length (aa)	473	247
Molecular weight (kDa)	53.07	26.66
Molar extinction coefficient	54530	13370
Isoelectric point (pI)	4.78	4.15

Test expression

To determine an optimal expression procedure, small-scale test expressions were carried out under different conditions. Samples were taken at different time points, and whole protein cell extracts as well as soluble and insoluble protein samples were monitored by SDS-PAGE. Fig. 4-33 depicts the expression of GST-InlG' in *E. coli* BL21 CodonPlus cells during an o.n. expression at 22 and 37°C. The fusion protein with a MW of 53 kDa runs slightly higher in the SDS-Gel, namely at ~60 kDa, a typical phenomenon of internalin-family proteins.

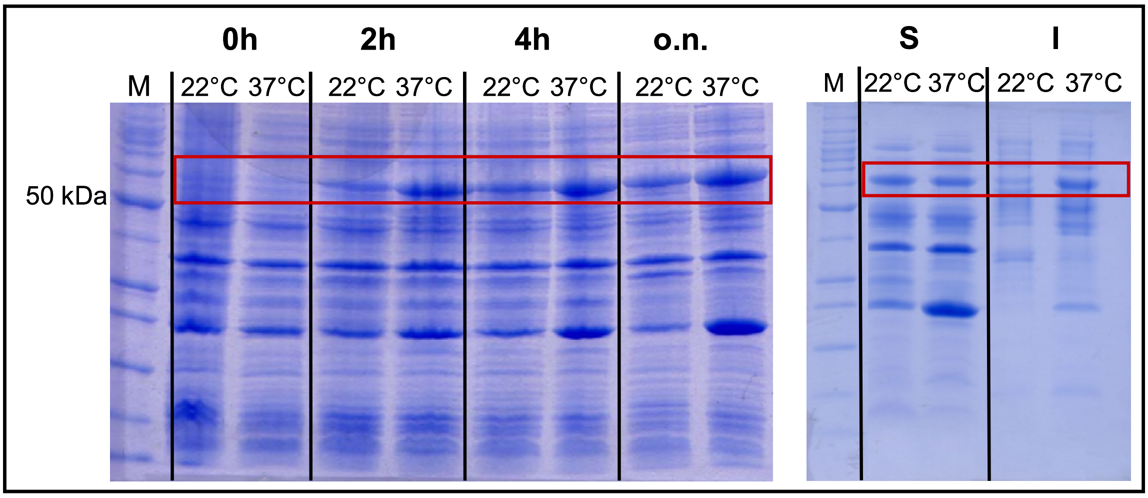


Fig. 4-33: Test expression of GST-InlG' in *E. coli* BL21 CodonPlus cells. Left: Cell extract after 0, 2 and 4 h and overnight (o.n.) at 22 and 37°C. Right: Soluble (S) and insoluble (I) cytoplasmic proteins after o.n. expression at 22 and 37°C. Red boxes: bands at ~60 kDa, containing GST-InlG'. M: Marker. Note that the higher expression level at 37°C is lost to inclusion bodies, since the soluble fractions from 22 and 37°C contain roughly equal amounts of GST-InlG'. Expression at 22°C thus yields similar amounts of soluble protein but under milder conditions.

Preparative expression and purification

InlG' was purified from *E. coli* BL21 CodonPlus grown in 2 L LB-medium as described in section 3.4.2. The protein was produced at 22°C (see expression test above). As depicted in Fig. 4-34, lane 'S', a substantial amount of soluble fusion protein was produced. Binding to the glutathione-sepharose resin was efficient, as seen by the lack of fusion protein bands in the flow-through ('FT') and wash samples. However, small amounts of protein leaked out from the resin upon extensive washing with PBS. Elution with 10 mM reduced glutathione was successful, but due to high yields, a fourth elution step was necessary to harvest the residual protein from the resin, as seen in the respective sample ('Resin'). Proteolytic cleavage of the eluted fusion protein was complete after digestion o.n. with PreScission® protease, and no residual fusion protein was visible on the SDS-Gel ('D').

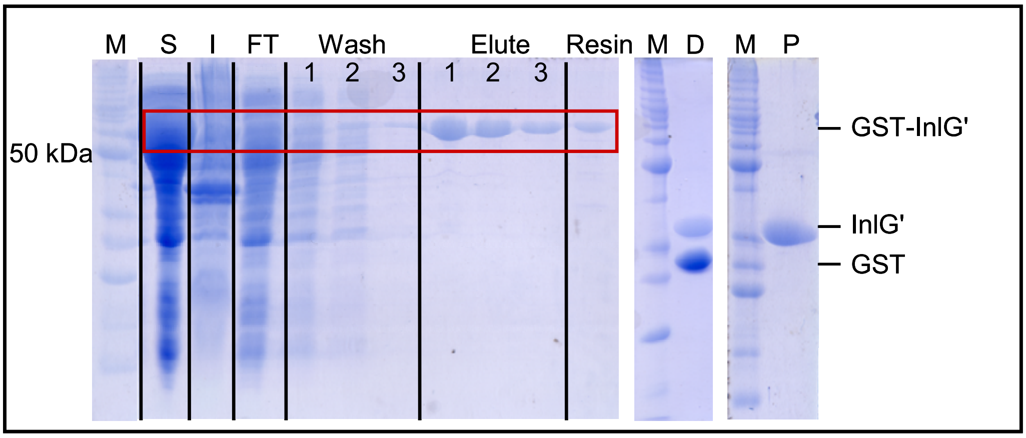


Fig. 4-34: Production, affinity chromatography and proteolytic digest of GST-InlG' in *E. coli* BL21 CodonPlus cells. M: Marker, S: soluble fraction, I: insoluble fraction, FT: flow-through, D: PreScission® digest (19°C, o.n.), P: Pure InlG' in 10 mM HEPES, pH 7.5, used for crystallization.

Following re-chromatography and IEC with a gradient of 100 – 400 mM NaCl in 10 mM HEPES, pH 7.5, InlG' was sufficiently pure for crystallization ('P'). The protein solution was desalted by dialysis and concentrated to 9-13 mg/mL. The overall yield of pure InlG' was ~25 mg per L of LB-medium.

Before crystallization setups, the concentrated solution of purified InlG' in 10 mM HEPES, pH 7.5, was analysed by N-terminal sequencing and DLS in order to check for N-terminal intactness, aggregation and heterogeneity. The hydrodynamic radius was measured to be 2.2 nm, leading to a calculated MW of ~20 kDa, with a proportionate polydispersity of 22 % (Fig. 4-35).

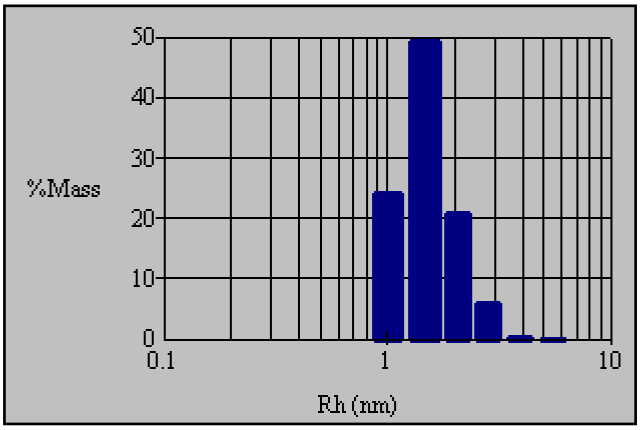


Fig. 4-35: Regularization histogram of purified InlG'. The presence of only one relatively sharp peak around 2.2 nm indicates the absence of high-molecular aggregates.

4.3.2 Crystallization

Sitting drop crystallization setups of InlG' were prepared with the following fast screens: JBS 1-10, Hampton Crystal I + II, Crystal Lite, PEG/Ion, Grid PEG 6000, Cryo and Grid PEG/LiCl. Eight conditions produced small crystals. Because these conditions were chemically partly similar, they were grouped to create three lines of optimization:

Table 4-8: Initial crystallization conditions for InlG'.

Line	Precipitant (buffer)	Type of crystals
1	PEG 4000 (Na acetate)	Large needles, 'combs'
2	$(NH_4)_2SO_4$	Needle tufts
3	Na citrate	Thin needles, 'barbells'

Various crystals of InlG' grown in hanging drop optimization trials are shown in Fig. 4-36.

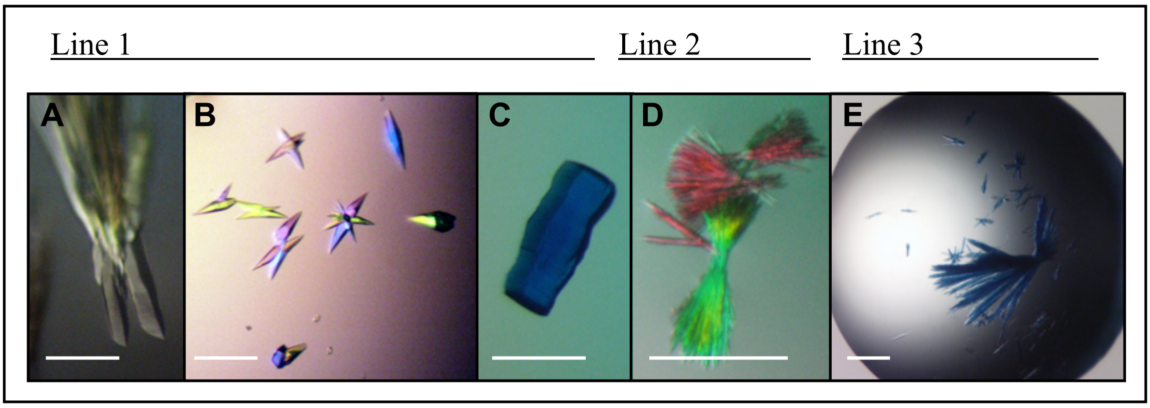


Fig. 4-36: Crystals of InlG'. (A) 31 % PEG 4000, 0.1 M Tris/HCl pH 8.3, 0.2 M Na acetate, 25°C. **(B)** 31 % PEG 4000, 0.1 M Tris/HCl pH 8.5, 0.2 M Na acetate, 20°C, streak seeding. **(C)** 30 % PEG 4000, 0.1 M Tris/HCl pH 8.5, 0.2 M Na acetate, 20°C. **(D)** 1.5 M Na citrate pH 7.5 20°C. **(E)** 2.2 M $(NH_4)_2SO_4$, pH 7, 20°C. Scale bar: ~0.1 mm.

X-ray diffraction data were eventually collected from crystals of optimization line 1.

4.3.3 X-ray data collection and structure solution

A needle-shaped crystal from the initial screen condition 1 (JBS2-H4: 30 % PEG 4000, 0.1 M Tris HCl pH 8.5, 0.2 M Na acetate) diffracted to a resolution of 2.05 Å. An X-ray diffraction data set of 244 images covering an oscillation angle of 122° was collected at the in-house rotating anode generator (Rigaku MSC). The structure was solved by molecular replacement (PHASER) using a model based on InlB (PDB code 1D0B). The solution was refined by simulated annealing (CNS) and continued with restrained refinement using REFMAC5. Data statistics and refinement details are listed in Table 4-9.

Table 4-9: Data collection and refinement statistics for InlG'.

Diffraction Data		**Refinement statistics**	
Space group	$C222_1$	Resolution range*	30 – 2.05
Unit cell dimensions a, b, c, Å	50.6, 92.7, 276.4	R_{work}/ $R_{free,}$ %	17.6 / 24.6
		Ramachandran plot¶	
Wavelength, Å	1.54	Preferred (%/#)	92 / 660
Resolution range, Å*	30 – 2.05 (2.12 -2.05)	Allowed (%/#)	8 / 57
Completeness, %*	98 (97)	r.m.s.d. bonds, Å/angles,	0,019/1.754
Unique reflections*	4524	# atoms protein/water	5547/431
I/σI*	11 (3.5)		
Redundancy*	4.2 (4.4)		
R-merge, %*	10.6 (45.8)		
Wilson B-factor, Å^2	21.0		

*Values in parentheses refer to the shell of highest resolution

¶ no residues in disallowed regions

4.3.4 InlG' adopts a typical internalin fold

In accordance with its close relationship to other internalin family members, the functional domain of InlG adopts a typical internalin fold. The cap and interrepeat domains do not exhibit significant differences compared to published internalin structures. With a LRR-region of only five repeats, the overall curvature of InlG is less pronounced than in other internalin structures, amounting to just 44° (see Table 4-10). Despite the marginal curvature, several hydrophobic side chains (Met102, Tyr126, Phe146 and Tyr168) face the solvent on the concave (β-) side of the InlG LRR region, as in InlJ (see above) and other internalin-family members, providing an extended interface to recognize a putative receptor protein.

The structure of InlG' supports its probable role in receptor binding / signalling. This agrees with knockout studies, that infer an accessory function for InlG as part of InlA-dependent host cell invasion (Bergmann *et al.*, 2002).

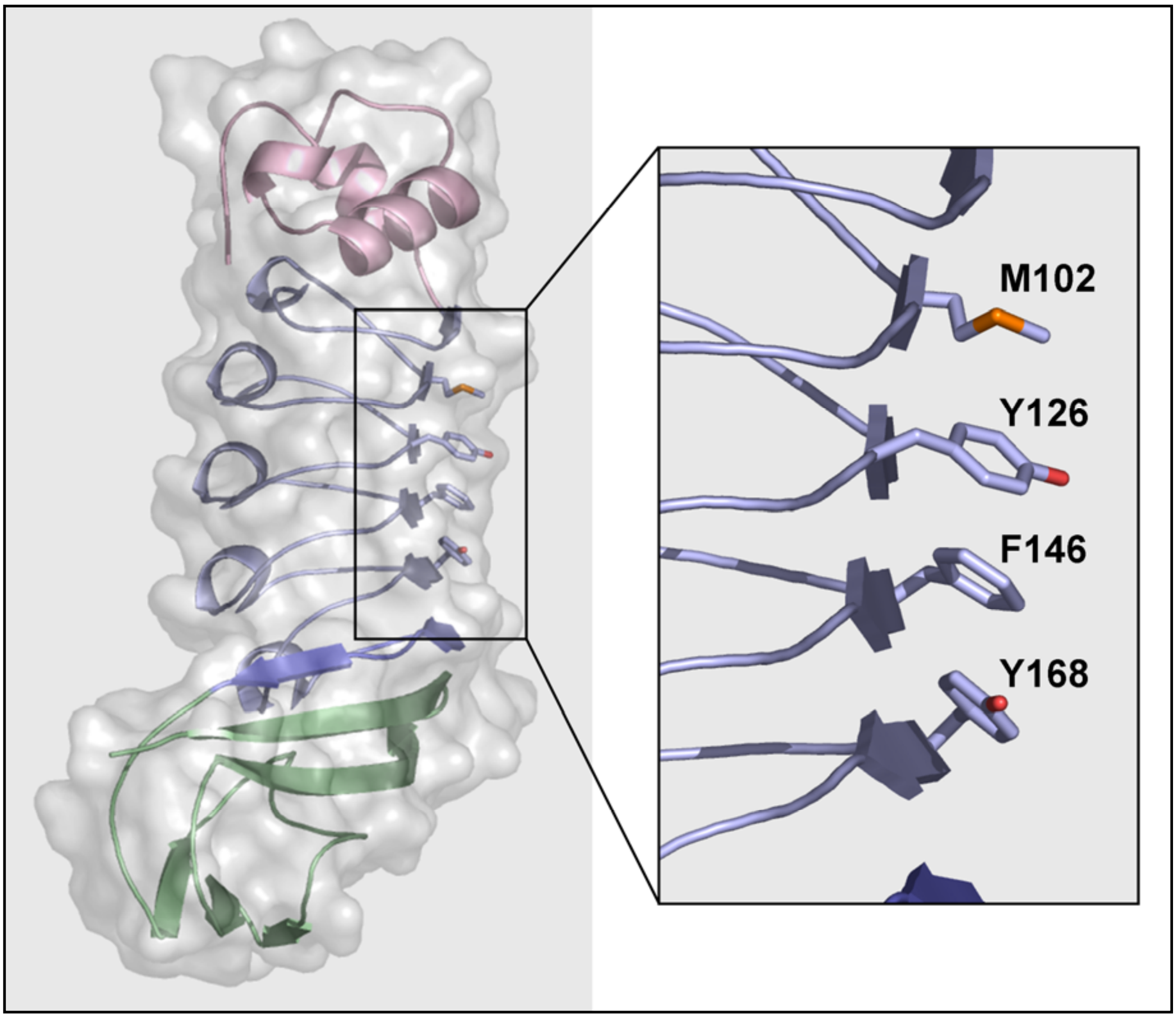


Fig. 4-37: Cartoon depiction of InlG', comprising the α-helical cap (light pink), five leucine-rich repeats (blue) and the Ig-like interrepeat domain (green). Solvent exposed hydrophobic residues are shown in stick representation and labelled in the zoom box (right).

4.4 Geometric analysis of LRR-proteins

The overall curvature and twist of LRR-protein structures have previously mostly been analysed in a qualitative fashion. Some standardized geometric analyses have been undertaken, especially to quantify the twist of the β-sheets as well as the radius and angle of protein curvature (Enkhbayar *et al.*, 2004; McEwan *et al.*, 2006). However, a more rigorous description of the LRR-geometry, taking into account the interdependence of the three rotation angles - curvature, twist and lateral bending - required to describe the overall geometry, has not been published to date.

To provide a consistent description of LRR protein geometry, the orientation of each repeat can be described with respect to the preceding repeat, allowing each angle to be calculated separately and independent of other geometric parameters. For this purpose, each LRR is represented by a defined triangle based on three invariant atoms per repeat (A, B and C in Fig. 4-38A and Fig. 4-29). Because the β-strand on the concave side of the protein is the most conserved structural feature of each repeat, the C_α-carbon atoms of residues at either end of the β-strand were chosen as points A and B (Pos. 2 and 4, Fig. 4-29). The outer, convex side of the LRR is more variable. Here, the C_α-atom of the conserved hydrophobic residue across from the β-sheet is chosen as atom C (pos. 15 in 21-residue repeats, pos. 16 in 20- or 22-residue repeats).

To simplify the geometric description, three mutually perpendicular planes are defined for each pair of LRRs: The basal, green plane (Fig. 4-38B) incorporates the reference triangle A_1-B_1-C_1 representing the first LRR of each LRR-pair. The lateral, pink plane is perpendicular to the basal plane (or A_1-B_1-C_1) and parallel to the vector A_1-B_1. The facial, blue plane is perpendicular to both A_1-B_1-C_1 and A_1-B_1.

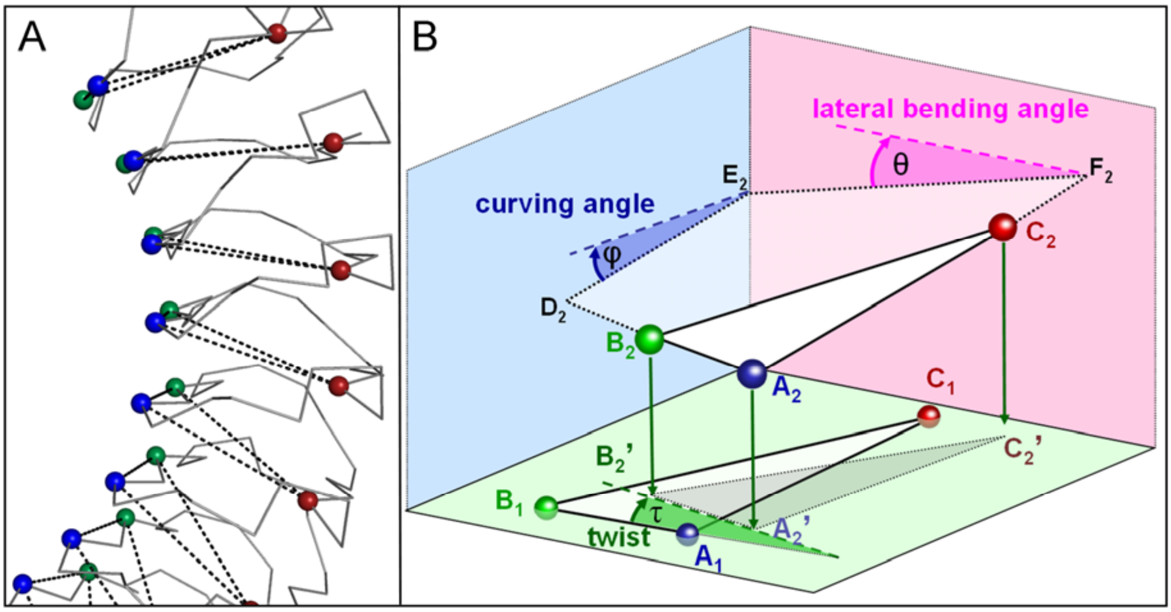


Fig. 4-38: Defining rotation angles for successive LRRs. (A) Each LRR is represented by a plane, using three invariant C_α-atoms defined in Fig. 4-29. **(B)** Triangles A_1-B_1-C_1 and A_2-B_2-C_2 represent consecutive LRRs. The base plane (light green) contains the first LRR. The twist angle τ (dark green) is the angle between A_1-B_1 and A_2'-B_2', the projection of A_2-B_2 onto the base plane. The facial plane (light blue) is perpendicular to both the base plane and to A_1-B_1. The curving angle φ (dark blue), defining the horseshoe-type bending of the protein, corresponds to the angle between the intersections of A_2-B_2-C_2 (line D_2-E_2) and A_1-B_1-C_1 with the facial plane. The lateral plane (light pink) is perpendicular to the base plane but parallel to A_1-B_1. The lateral bending angle θ (dark pink) corresponds to the angle between the intersections of A_2-B_2-C_2 (line E_2-F_2) and A_1-B_1-C_1 with the lateral plane.

Geometrically, the twist τ of a particular LRR (A_2-B_2-C_2) with respect to its predecessor (A_1-B_1-C_1) is thus simply the angle between A_1-B_1 and A_2'-B_2' - the projection of A_2-B_2 onto the basal plane (green angle, Fig. 4-38B). The curving angle φ of A_2-B_2-C_2 relative to A_1-B_1-C_1, causing the overall horseshoe shape of the LRR-domain, is the angle of the projection of A_2-B_2-C_2 in the facial (blue) plane relative to the basal plane (blue angle, Fig. 4-38B). Similarly, the lateral bending angle θ is the angle between the component of A_2-B_2-C_2 in the lateral (pink) plane relative to that of the basal plane (pink angle, Fig. 4-38B).

The geometric analysis of all available structures of listerial internalins, InlJ, InlG, InlA, InlB, InlC, InlE and InlH from *L. monocytogenes,* YopM from *Yersinia pseudotuberculosis*, and the porcine ribonuclease inhibitor (RI) is summarized in Fig. 4-39, and Table 4-10.

Curvature

With average curving angles ranging from 10.9° for InlJ to 13.5° for InlB, the curvature of internalin-family proteins and YopM is remarkably constant. Compared to the other family members, the terminal four repeats of InlJ represent somewhat of an anomaly in that the curving angles decrease from an average of 12° for LRR2-11 to an average of 8° for LRR12-15. This change in curving angle coincides with the interruption of the cysteine- and asparagine-ladders (Fig. 4-28), indicating that these ladders may contribute significantly to the wedge shape of the LRRs. With an average of 19.3°, the RI is significantly more curved than the internalins and YopM. The amphipathic α-helices on the convex side of RI cause the repeats to adopt a more pronounced wedge-shape, while 3_{10}-helices present in most internalins (except InlJ) induce smaller curving angles.

Surprisingly, though, non-helical stretches in short repeats such as InlJ and YopM give rise to similar curvatures as those of 3_{10}-helix bearing repeats. Clearly, additional effects such as the embedded 'strings' of water molecules (see above) may compensate for the lack of thickness of extended loops compared to 3_{10}-helices.

The zigzag shape of the curving angle plot for RI (Fig. 4-39A) provides another example of the fact that the secondary structure of the variable region alone does not dictate the LRR-curvature. Despite α-helical variable domains throughout, the curving angles alternate between an average of 27° and 12°, correlating with two distinct repeat lengths of 28 (A-type) and 29 residues (B-type). Differences in curving angles appear to be caused by distinct patterns of conserved residues in the two repeats, rather than the length of the repeat itself. A similar zigzagging in the curving angle plot is observed in YopM for LRR4-10, coinciding with alternating 20 and 22 residues per repeat in this part of the molecule.

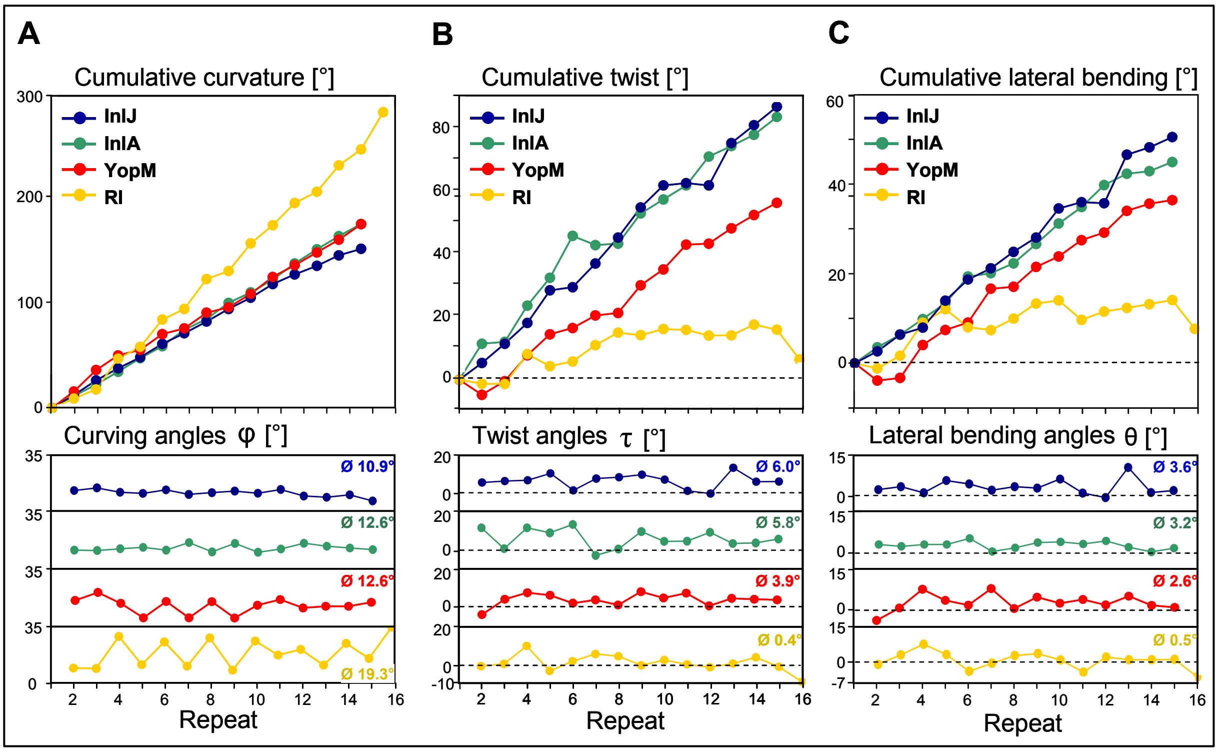


Fig. 4-39: Graphical depiction of the LRR-geometry. (A) Curvature, **(B)** twist, and **(C)** lateral bending for the LRR-proteins InlJ, InlA, YopM and the porcine ribonuclease inhibitor (RI). Cumulative values are provided in the top panel whereas values for each repeat are plotted in the lower panels. Twist angles are generally defined as positive if the associated rotation as defined in Fig. 4-38 is clockwise, negative if counter-clockwise. Lateral binding angles are defined as positive if the distance B_2-B_2' in Fig. 4-38 is smaller than A_2-A_2'.

Twist

The surface-exposed β-sheets of InlJ, InlA and YopM generally exhibit a clockwise (positive) twist (Fig. 4-39B). Average twist angles per repeat, however, vary substantially from one protein to the next. For InlJ, the cumulative twist angle is 85° over 15 repeats or 6.0° per repeat. This is very similar to InlA, where an identical number of repeats result in an overall twist of 82°, amounting to a twist of 5.8° per repeat. In other members of the internalin family, the average twist ranges from as low as 2.5° for InlC to 6.4° for InlB. With an average twist of 3.9° YopM falls into the internalin range. The porcine RI deviates significantly from the other LRR-proteins investigated in that its cumulative twist amounts to only 6° for 16 repeats, or an average of 0.4° per repeat – resulting in its planar horseshoe shape.

Examining the twist values in more detail reveals that the twist is not strictly unidirectional. In InlJ, for example, β-strand 12 is rotated -0.7° with respect to that of LRR11. LRR6 with a twist of 0.95°, similarly, falls short of the 6.0° average twist of this protein. In both LRR6 and LRR12 the cause for this deviation appears to be the insertion of a hydrophilic residue after position 8 in the post-β-strand loop (Fig. 4-29C, Fig. 4-39B). The inserted residues, Lys211 in

LRR6 and Lys338 in LRR12, partially mimic residue 9 from adjacent repeats, but the additional residue disrupts the repeat stacking pattern inducing an anomalously small twist angle. The constrained neighbourhood forces the Lys210 backbone into an unconventional conformation, making it an outlier in the Ramachandran plot.

LRR7 of InlA is similarly rotated by -2.8° relative to LRR6, which itself has the strongest clockwise twist of 13° relative to LRR5 resulting in a distinct kink in the twist graph (Fig. 4-39B). This behaviour appears to be caused by the deletion of a residue in LRR6 and the concomitant disruption of the LRR geometry. The large buried side-chain of Phe193 in LRR6 furthermore displaces Ile214 in LRR7 away from the hydrophobic core, contributing to the negative twist angle. Phe105 in LRR2 has a similar effect on Met127 in LRR3, giving rise to another twist angle (0.56°) significantly below the average of 5.8°. Inserting a 22nd residue in LRR6 (Wollert *et al.*, 2007a) reduces but does not abolish the 'kink'. The twist is thus not primarily dictated by the length of the repeat but rather by the constituent residues and their ability to integrate amongst the residues of neighbouring repeats.

Table 4-10: Geometric analysis of LRR-architecture.

Parameter	**InlA**	**InlB**	**InlC**	**InlE**	**InlG**	**InlH**	**InlJ**	**YopM**	**RI**
PDB code	1O6V	1H6T	1XEU	1XYZ	n.d.	1H6U	n.d.	1JL5	2BNH
No. of repeats	15	7	6	8	5	8	15	15	16
Repeat length	22	22	22	22	22	22	21	20	28/29
Variable region	3_{10}	3_{10}	3_{10}	3_{10}	3_{10}	3_{10}	ext.	ext.	α
Residues for A,B,C	2,4,16	2,4,16	2,4,16	2,4,16	2,4,16	2,4,16	2,4,15	2,4,16	2,4,22
Curvature (°):									
avg. per repeat	12.6	13.5	11.2	12.7	11.0	12.3	10.9	12.6	19.3*
cumulative	176.1	80.7	55.9	88.7	43.9	86.0	152.2	175.7	289.3
Twist (°):									
avg. per repeat	5.8	6.4	2.5	3.8	6.2	4.6	6.0	3.9	0.4
cumulative	81.5	38.5	12.6	26.4	24.6	32.2	84.7	54.9	6.3
Lateral bending (°):									
avg. per repeat	3.2	2.7	0.8	2.5	3.0	2.3	3.6	2.6	0.5
cumulative	45.0	16.0	3.9	17.8	12.2	16.4	50.5	36.4	7.7

*A:11.6, B:26.0; n.d.: not deposited

The negative twist of –4.7° of LRR2 in YopM relative to LRR1 appears to be caused by the shape of the preceding cap domain. Though α-helical like the internalin cap, the cap of YopM is structurally distinct, distorting the β-strand of LRR1.

In terms of individual twist angles, the RI deviates dramatically from all internalins and YopM. In RI, eight of fifteen repeats (LRR2, 3, 5, 9, 11, 12, 15 and 16) have negative twist angles. In general these twist angles are rather small (less than -1°) except for LRR5 (-3.9°) and LRR15 (-9.2°). Of the LRRs with positive twist angles, three (LRR6, 10, 13) are smaller than 2°, while the largest is 9.3° (LRR4). If the last two repeats of RI are ignored, the cumulative twist for 14 repeats is 15.5°, or just over 1° per repeat. This slight positive twist is, however, eliminated by the -9.2° twist of LRR15 (Fig. 4-39B).

Lateral bending

The described calculations indicate considerable cumulative lateral bending angles in all internalins and YopM, indicating that in general C-terminal ends of β-strands are more closely spaced than corresponding N-terminal ends (Fig. 4-39C).

As for the other angles, irregularities in lateral bending angles are indicative of distinct structural features. For example, in LRR11 of InlJ, the bulky side-chain of Ile313 C-terminal of the β-strand that faces outward appears to function as a non-polar spacer with respect to hydrophilic Ser292 located at the equivalent position in LRR10. This results in a bending angle below average value. Similarly, in LRR12, the regular placement of cysteine within the Cys-ladder at position 5 is interrupted. This is replaced by a cysteine at position 7 that faces the preceding repeat within the hydrophobic core of the protein. This misalignment reduces the lateral bending by preventing the C-terminal ends of the β-strands from packing as tightly compared to neighbouring repeats. A large bending angle in the following LRR13 is caused by the cysteine in position 5 of LRR12 being replaced by alanine. The smaller residue allows the repeats to approach each other more closely at the C-terminal end of the β-sheet.

4.5 GFP-labelling system for internalins

The biological function of the majority of internalin family members encoded by *L. monocytogenes* is currently not clear. To systematically analyse the interaction of these proteins with eukaryotic cells *in vitro*, a uniform set of internalin-GFP fusion constructs was established. Each internalin gene was cloned into three target vectors: *pETM30GFPfus* for the production of GST-internalin-GFP fusion proteins in *E. coli*, and *pEGFP-C1* and *pEGFP-N3* for heterologous expression of respectively GFP-internalin and internalin-GFP fusion constructs in eukaryotic cells.

4.5.1 Cloning strategies

Due to different available constructs of the genes for InlA, InlB, InlC, InlE, InlF, InlG and InlH, individual cloning strategies were developed. Table 4-11 provides an overview describing previously available constructs and resulting cloning procedures used to produce stereotype GFP-fusion constructs of all available internalin genes. The oligonucleotide numbering refers to Table 3-5.

Table 4-11: Cloning strategies for the production of stereotype GFP-InlX-fusion constructs.

Donor Plasmid (Source)	**Production of insert gene**	**Restriction sites in insert**	**Recipient Plasmid**	**Restriction sites in recipient**
InlA$_{36-496}$				
pGEX-6p-1-InlA$_{36-496}$ (T.Hain)	PCR (oligos no. 2, 3)	EcoRI, BamHI	*pEGFP-C1*	EcoRI, BamHI
			pEGFP-N3	
	PCR (oligos no. 1, 2)	NcoI, BamHI	*pETM30GFPfus*	NcoI, BamHI
InlB$_{36-321}$				
pGEX-6p-1-InlB$_{36-321}$ (T.Hain)	PCR (oligos no. 5, 6)	HindIII, BamHI	*pEGFP-C1*	HindIII, BamHI
			pEGFP-N3	
	PCR (oligos no. 4, 5)	NcoI, BamHI	*pETM30GFPfus*	NcoI, BamHI
InlE$_{36-342}$				
pGEX-2TK-InlE$_{1-342}$ (G.Göbel)	PCR (oligos no. 8, 9)	EcoRI, BamHI	*pEGFP-C1*	EcoRI, BamHI
			pEGFP-N3	
	PCR (oligos no. 7, 8)	NcoI, BamHI	*pETM30GFPfus*	NcoI, BamHI
InlF$_{36-795}$				
pGEX-6p-1-InlF$_{36-795}$ (J.Ehinger)	Restriction of donor	BamHI, EcoRI	*pEGFP-C1*	BglII, EcoRI
	PCR (oligos no. 11, 12)	EcoRI, BamHI	*pEGFP-N3*	EcoRI, BamHI
	PCR (oligos no. 10, 11)	NcoI, BamHI	*pETM30GFPfus*	NcoI, BamHI
InlF$_{36-454}$				
pGEX-6p-1-InlF$_{36-454}$ (J.Ehinger)	Restriction of donor	BamHI, EcoRI	*pEGFP-C1*	BglII, EcoRI
	PCR (oligos no. 12, 13)	EcoRI, BamHI	*pEGFP-N3*	EcoRI, BamHI
	PCR (oligos no. 10, 13)	NcoI, BamHI	*pETM30GFPfus*	NcoI, BamHI
InlG$_{36-460}$				
pGEX-6p-1-InlG$_{36-460}$ (J.Ehinger)	PCR (oligos no. 16, 17)	EcoRI, KpnI	*pEGFP-C1*	EcoRI, KpnI
			pEGFP-N3	
	PCR (oligos no. 14, 15)	NcoI, NheI	*pETM30GFPfus*	NcoI, NheI

InlG$_{36-277}$				
pGEX-6p-1-InlG$_{36-277}$ (J.Ehinger)	PCR (oligos no. 16, 19)	EcoRI, KpnI	*pEGFP-C1*	EcoRI, KpnI
			pEGFP-N3	
	PCR (oligos no. 14, 18)	NcoI, NheI	*pETM30GFPfus*	NcoI, NheI
InlH$_{36-519}$				
pGEX-2TK-InlH$_{1-548}$ (G.Göbel)	PCR (oligos no. 20, 22)	EcoRI, BamHI	*pEGFP-C1*	EcoRI, BamHI
			pEGFP-N3	
	PCR (oligos no. 20, 21)	NcoI, BamHI	*pETM30GFPfus*	NcoI, BamHI

The genes were amplified by PCR, introducing restriction enzyme cleavage sites in tailored primer overhangs. The only exception was the cloning of InlF into *pEGFP-C1*, where the gene could be cleaved out of the donor plasmid and directly ligated into linearised target vector, because of compatible cohesive ends.

Successful cloning into the *pETM30GFPfus* system was verified by IPTG-inducible fluorescence (Fig. 4-40) of transformed *E. coli* BL21 CodonPlus cells and DNA sequencing.

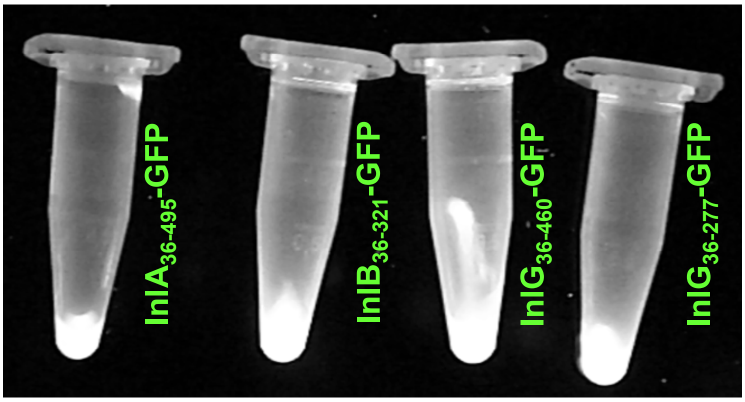


Fig. 4-40: *E. coli* BL21 CodonPlus cells expressing GFP-Inl-fusion proteins from the vector *pETM30GFPfus*.

The *pEGFP-C1* and *–N3* constructs were controlled by test restriction and DNA sequencing. All constructs (except for InlE$_{36-342}$ in *pETM30GFPfus*) were successfully cloned into the three vectors. The complete set of constructs provides the basis for systematic studies to directly compare the interactions of single internalin proteins with eukaryotic cells *in vitro*.

4.5.2 Intracellular localization of GFP-labelled internalins

This experiment was aimed at revealing possible subcellular localization of the GFP-tagged internalins and to monitor associated changes in cell morphology, growth and cytoskeletal organization. Transfected A-431 cells are depicted in Fig. 4-41. The set of internalin constructs, each N- (*pEGFP-C1*) and C-terminally (*pEGFP-N3*) tagged with GFP, were transfected into the human epithelial carcinoma cell line A-431.

In all samples, fluorescence is uniformly distributed throughout the cytoplasm and in the nucleus of transfected cells, regardless of the terminus to which GFP was fused. Similarly, the presence or absence of GW-domains in the constructs of InlF and InlG did not alter the subcellular distribution. However, expression levels were low throughout, and decreased with increasing construct length. Given intact proteins, nuclear localization could only be caused by active transport. Furthermore, a slightly higher rate of apoptosis occurred in transfected cells. A distinct or exclusive recruitment of any internalin protein to a special subcellular site, e.g. cell-cell-contacts, nucleus or plasma membrane was not found. This would confirm the hypothesis that internalins bind extracellular receptors, rather than functioning as intracellular effectors.

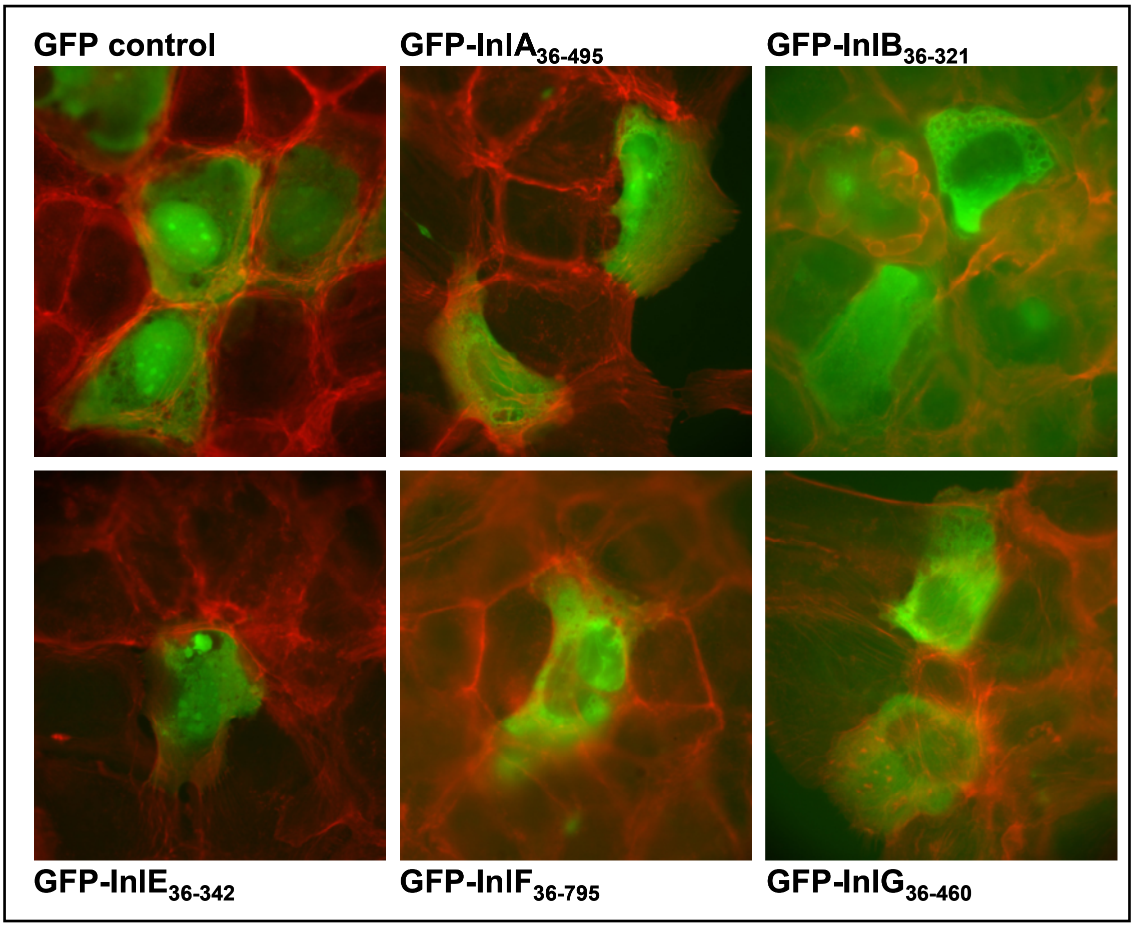


Fig. 4-41: Subcellular distribution of GFP-labelled internalin proteins heterologously expressed in A-431 cells. Intracellular distributions of fusion proteins are similarly uniform, and transfected cells do not show any altered morphologies, except for a generally higher rate of apoptosis as compared to the GFP-transfected control.

5 Discussion

5.1 The autolysin Auto

5.1.1 A lysozyme-like fold with *N*-acetylglucosaminidase activity

Auto is a member of the glycoside hydrolase family GH73 (CAZY database, Coutinho & Henrissat, 1999), a family that currently includes 617 bacterial and 33 viral enzymes. Prior to this study, no information on the structure and catalytic mechanism of any family member was available. GH73 family members with high sequence similarity to the catalytic domain of Auto have been demonstrated to have *N*-acetylglucosaminidase activity on PG, consistent with this study. These enzymes include AtlA from *Enterococcus faecalis* (53 % sequence identity), AcmA (51 %) and AcmB (40 %) from *Lactococcus lactis*, Mur-2 from *Enterococcus hirae* (50.6 %) and LytG from *Bacillus subtilis* (50 %) (Eckert *et al*., 2006; Eckert *et al*., 2007; Horsburgh *et al*., 2003; Huard *et al*., 2003; Steen *et al*., 2005).

It is thus particularly interesting that Auto is structurally related to eukaryotic g-type lysozyme, the soluble lytic transglycosylase Slt70 from *E. coli* (both GH23) and family GH19 chitinases. GH23 enzymes preferentially cleave the NAM-(1,4)-NAG bond, chitinases the NAG-(1,4)-NAG bond, and Auto, by cleaving the NAG-(1,4)-NAM bond now represents the third type of enzymatic function accomplished by a related lysozyme-like fold. Clearly, the fold has co-evolved in bacteria and eukaryotes as a versatile scaffold for PG hydrolysis at different positions and with differing products.

Conserved structural elements between the classes GH73, GH23 and GH19 are the overall bilobal architecture, the helical core of the enzyme, and the ‘catalytic’ glutamate in a hydrophobic environment (Fig. 4-21 and Fig. 5-1). However, a striking difference between Auto and all related structured is the less confined substrate binding groove (Fig. 5-1).

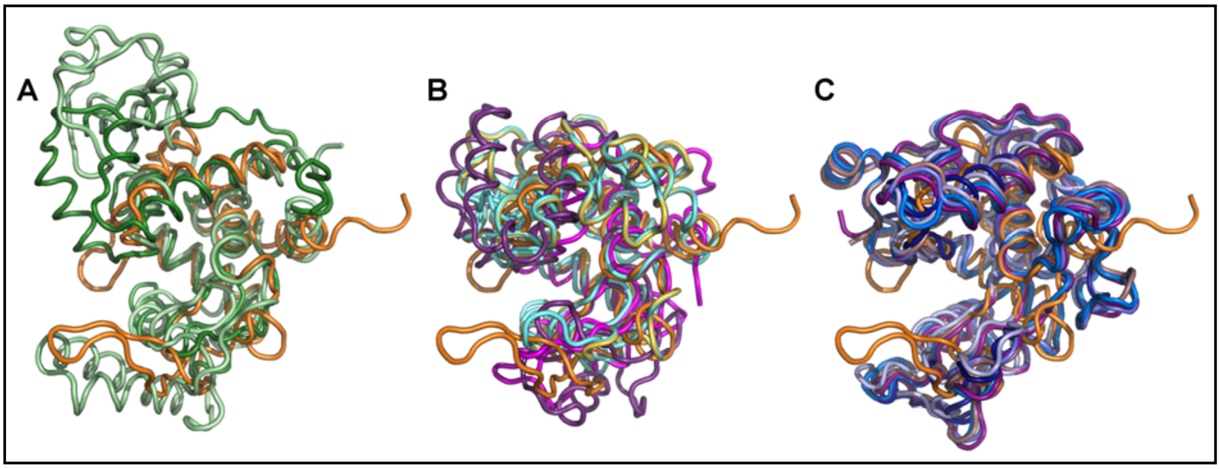


Fig. 5-1: Three enzyme classes with a lysozyme fold. Overlay of Auto (orange, inhibitory α0 omitted) with **(A)** lytic transglycosylases, **(B)** c-, p- and g-type lysozymes and **(C)** class GH19 chitinases. The less confined substrate binding groove in Auto stands out.

5.1.2 Implications for the catalytic mechanism

Comparing 458 sequences of the family GH73 clearly demonstrates that the substrate binding groove and the immediate vicinity of the catalytic glutamate are the most conserved regions of the enzyme (Fig. 5-2). Similar to g-type lysozymes, Slt70 and GH19 chitinases, Auto lacks a closely spaced second catalytic carboxylate, as provided by Asp52 in HEWL. Therefore, a retaining double-displacement mechanism involving a covalent glycosyl-enzyme intermediate or stabilization of an oxocarbenium ion intermediate can be ruled out for Auto, as both mechanisms would require a second carboxylate in immediate vicinity to the NAG C-1 atom. In phage T4 lysozyme, the slightly more distant Asp20 acts as a general base assisting a nucleophilic water attack from the opposite side of the sugar ring (Kuroki *et al.*, 1999). Such a single-displacement mechanism is also thought to hold true for family GH19 chitinases and probably also GEWL (Fukamizo, 2000). This mechanism implies inversion of the anomeric configuration and, more importantly, tolerates larger distances between the two carboxylates than in HEWL (McCarter & Withers, 1994). In Auto, Glu156, located within the β-hairpin at a distance of 13 Å from Glu122, could provide a second catalytic carboxylate for assisting a water nucleophilic attack from the α-side of the substrate. Yet, to involve this residue in catalysis, the enzyme would need to undergo a significant conformational rearrangement upon substrate binding, comparable to the movement of the β-sheet towards the ligand that has already been reported for c-Slt70 upon binding of the inhibitor bulgecin (Thunnissen *et al.*, 1995b). The fact that an acidic residue at the position of Glu156 is conserved in the vast majority of available sequences from family GH73 supports its role in catalysis (Fig. 5-2).

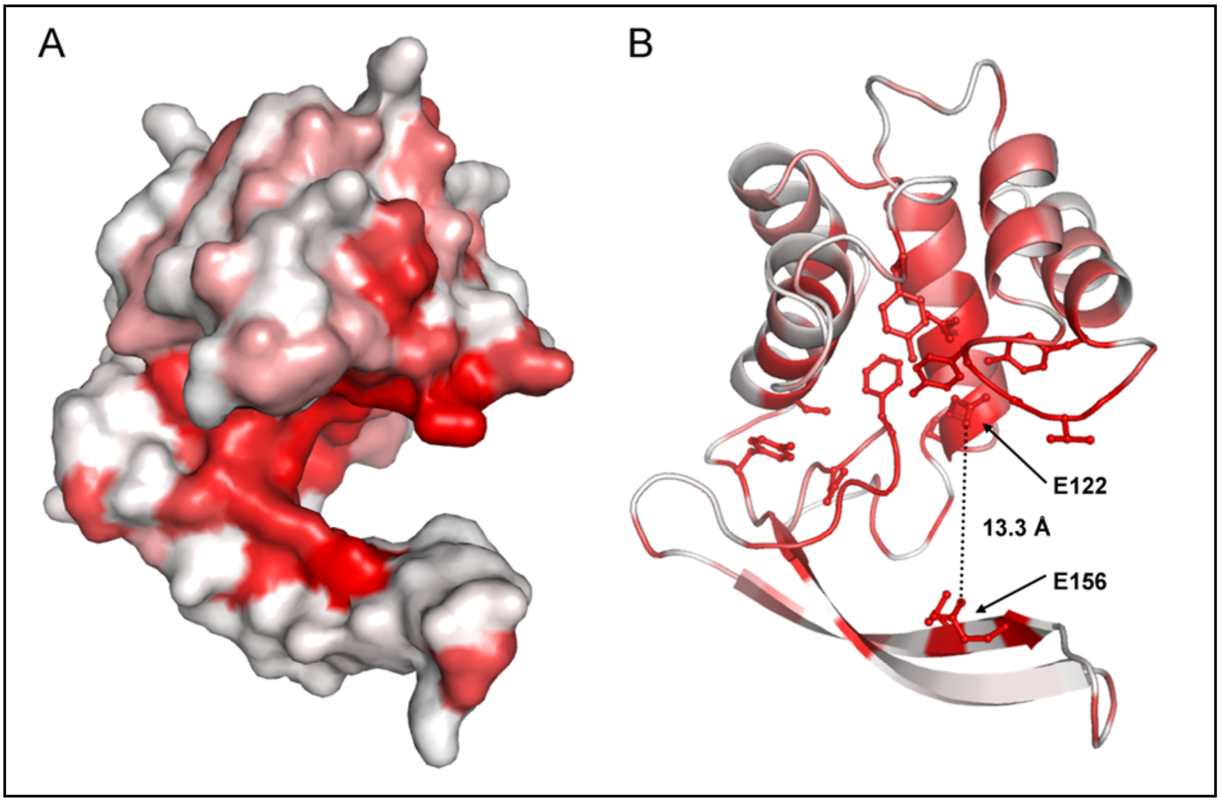


Fig. 5-2: Residue conservation in Auto$_{act}$. The degree of conservation, as obtained from an alignment of Auto$_{act}$ with 457 other GH73 sequences, was colour-coded and mapped onto the surface of the crystal structure of pro-Auto (inhibitory α0 omitted). Red = conserved, pink = partly conserved, white = not conserved. **(A)** Surface representation. The substrate binding groove has the highest degree of conservation. **(B)** Cartoon representation. Conserved residues are shown in ball-and-stick. The 'catalytic' Glu122 and the putative second catalytic Glu156 are 13.3 Å apart.

There is, however, also evidence that a single acidic group may be sufficient for catalysis, as also discussed for GEWL and Slt70 (Thunnissen et al., 1994; Weaver et al., 1995). To substitute for an appropriately positioned carboxylate residue, functional groups of the substrate could be involved. For Slt70 and the plant chitinase hevamine a mechanism of substrate-assisted catalysis has been proposed, with a nucleophile being provided from the C-2-N-acetamido group of the substrate (Terwisscha van Scheltinga et al., 1995; Thunnissen et al., 1995b). For a HEWL D52A variant and GEWL, proposed mechanisms involve a carboxylate from the peptide bridge in PG to stabilize a putative oxocarbenium ion intermediate (Matsumura & Kirsch, 1996). This model of substrate-assisted catalysis mediated by a peptide carboxyl is consistent with the finding that Auto$_{act,}$ like Slt70 and GEWL, elicits higher activity on peptide-substituted cell walls or intact PG (Arnheim et al., 1973; Schindler et al., 1977; Thunnissen et al., 1995a) than on (NAG)$_x$, which is not cleaved.

5.1.3 Which PG fragment binds to Auto_{act}?

HEWL, which cleaves $(NAG)_6$, is efficiently inhibited by $(NAG)_3$ (Chipman & Sharon, 1969). Auto, however, does not cleave NAG-oligomers, while small components such as $(NAG)_3$, $(NAG)_4$ or NAG-NAM-dipeptide do not competitively inhibit degradation of *M. lysodeiktikus* PG by Auto. Auto thus appears to have a very low affinity for $(NAG)_x$ and probably only recognizes and binds NAM-containing oligosaccharides and/or larger components of the cell wall, potentially including the peptide bridge. The extended β-hairpin and the loop flanking the substrate binding groove expose several aromatic residues to the solvent, allowing for the binding of larger PG substrate molecules. Since Auto_{act} exhibits *N*-acetylglucosaminidase activity with a lysozyme-like (i.e. *N*-acetylmuramidase) fold, the carbohydrate binding subsites would need to differ from those in lysozymes. Most importantly, the sugar residues would need to shift by one subsite, to cleave the bond between a NAG-residue in subsite -1 and NAM in subsite +1. Correspondingly, the majority of residues involved in substrate-binding in lysozymes and c-Slt70 are not conserved in Auto. To describe the precise mode of saccharide ligand coordination within the binding groove would require a crystal structure of ligand-bound Auto_{act}. Crystallizing Auto_{act} alone or in complex with a substrate analogue did, however, not succeed.

5.1.4 Control of autolytic activity

Eukaryotic lysozymes and chitinases act as bactericides and fungicides that are usually harmless to the host. Autolysins, produced by the bacteria themselves, must be spatially and temporally tightly regulated to prevent uncontrolled cell wall degradation and hence cell lysis. The control of autolysin activity is generally not well understood. Proposed mechanisms include regulation of transcription, production of inactive precursors, modulation of PG composition and binding to restricted regions on the cell surface (Höltje & Tomasz, 1975; Smith *et al.*, 2000). The composition of lipoteichoic acids and teichoic acids in the cell wall are known to affect autolysin activity (Höltje & Tomasz, 1975).

Post-translational proteolytic cleavage is a simple mechanism to keep Auto in the “off” mode, until its activity is required. This mechanism is common to numerous proteolytic enzymes, including those of the digestive tract (Whitcomb & Lowe, 2007). Other autolysins, such as Atl from *S. aureus* (Foster, 1995) or AcmA from *L. lactis* (Buist *et al.*, 1995) have been reported to be processed from latent pro-enzymes to smaller active forms. The prophage endolysin PlyL of *Bacillus anthracis* is apparently allosterically inhibited by its C-terminal cell-wall binding domain until it encounters its substrate (Low *et al.*, 2005). This study now provides the first structural insight into such an autoinhibitory mechanism in which an N-terminal α-helical plug occupies the substrate binding groove and efficiently blocks the 'catalytic' glutamate.

5.1.5 Controlled Auto activity and listerial virulence

Activation of Auto

The sequence of the autoinhibitory region and the activation loop are unique to Auto and – more importantly – restricted to the pathogenic species *L. monocytogenes*, implying that its mode of activation is closely linked to the involvement of Auto in pathogenesis. Possibly, an environmental trigger, such as extracellular host proteases may be required to activate Auto. Candidate host proteases include cathepsin D, shown to accumulate in phagosomes of LoVo cells containing InlA- and InlB-coated latex beads, 30 minutes after bead uptake (Pizarro-Cerda *et al*., 2002). Another phagosomal cysteine protease was found to activate the broad range phospholipase PC-PLC involved in phagosomal escape (Marquis *et al*., 1997). Alternatively, endogenous listerial proteases such as the metalloprotease Mpl could activate pro-Auto in a phagosomal environment (Marquis *et al*., 1995).

Phagosomal escape

Especially the phagosomal step of infection may depend on autolytic activity. Efficient phagosomal escape requires the secretion of the proteins listeriolysin (LLO) and the phospholipases PI-PLC (PlcA) and PC-PLC (PlcB) (Portnoy *et al*., 1992; Smith *et al*., 1995). The rapid release and activation of these effectors, triggered by vacuolar acidification, accelerates the bacterial escape into the cytosol before phagolysosomal fusion can occur. The Gram-positive PG is known to act as a diffusion barrier, and the permeability of the cell wall is substantially decreased in *B. subtilis* and *S. pneumoniae* autolysin mutants (Dijkstra & Keck, 1996). Secreted staphylococcal enterotoxin B was similarly observed to associate with the cell wall and can only be extracted by treating bacteria with a cell wall hydrolase (Tweten & Iandolo, 1983). More specifically, rapid translocation of accumulated broad-range phospholipase C (PC-PLC) from the cell-membrane/cell wall interface across the PG is dependent on a decrease in pH (Snyder & Marquis, 2003). Auto, if activated in this step of infection by a phagosomal or endogenous protease, would be highly active as the vacuole acidifies (de Chastellier & Berche, 1994) due to its acidic pH-optimum, and could assist the rapid translocation of previously retained virulence factors. The idea of environmental conditions in the vacuole influencing autolysin activities and, concomitantly, protein translocation across the cell wall has been proposed before (Snyder & Marquis, 2003). This hypothesis implies a host-compartment dependent activation of Auto and would also account for the dramatic loss of invasiveness found for a *L. monocytogenes Δaut* mutant *in vitro* (Cabanes *et al*., 2004). Attachment of Auto to the cell wall through its GW-domains (Cabanes *et al*., 2004), allows for local puncturing of the PG, rather than complete lysis by freely diffusing Auto, allowing fast and controlled release of previously retained factors without losing overall cell wall integrity.

Steering of Auto by PG-deacetylation

The enzyme PgdA partly *N*-deacetylates PG mediating resistance to lysozyme and supporting listerial virulence (Boneca *et al*., 2007). Phagosomal escape is correspondingly impaired in a *L. monocytogenes pgdA*-mutant, linking bacterial escape mechanisms to the *N*-deacetylation of PG. In addition to host innate defence evasion, bacteria use PG *N*-deacetylation to regulate the activity of endogenous autolysins, as demonstrated for the GH73 family member AcmA from *L. lactis* (Meyrand *et al*., 2007). Our MS analyses confirm that listerial PG contains *N*-deacetylated glucosamine, and that it preponderates in incompletely digested murein fragments. Deacetylation thus appears to represent a steering mechanism of endogenous autolysin activity. Unmodified PG, by contrast, is prone to rapid degradation by activated Auto, resulting in uncontrolled autolysis. Excessive autolysis is thus presumably responsible for the loss of invasiveness of an *aut* overexpressing *L. monocytogenes* mutant (Cabanes *et al*., 2004), underlining the importance of tight regulation.

Fig. 5-3 depicts a simplified model of involvement of Auto in phagosomal escape.

Auto-induced PG turnover

Listerial uptake and phagosomal escape are known to be dependent on elevated PG turnover rates. How this recycling is achieved is not understood. In this regard it is interesting that a variant of InlA with a 6700-fold increased affinity for its receptor E-cadherin (Wollert *et al*., 2007b) does not ‘trap’ the corresponding *L. monocytogenes* mutant Lmo-InlA^{m} at the phagosomal membrane. Because both InlA and E-cadherin are physically bound to their respective cell-surfaces, combined with the fact that the high complex affinity is such that dissociation is very slow, the bacterial cell wall must rapidly turn over within the phagosome to prevent ‘trapping’.

In conjunction with previous data on the involvement of Auto in listerial infection, the molecular characteristics of the enzyme revealed in this work (facile activation by proteolytic cleavage, acidic pH-optimum, preference towards *N*-acetylated PG), support the interpretation that Auto is involved in the coordinated phagosomal escape of *L. monocytogenes*.

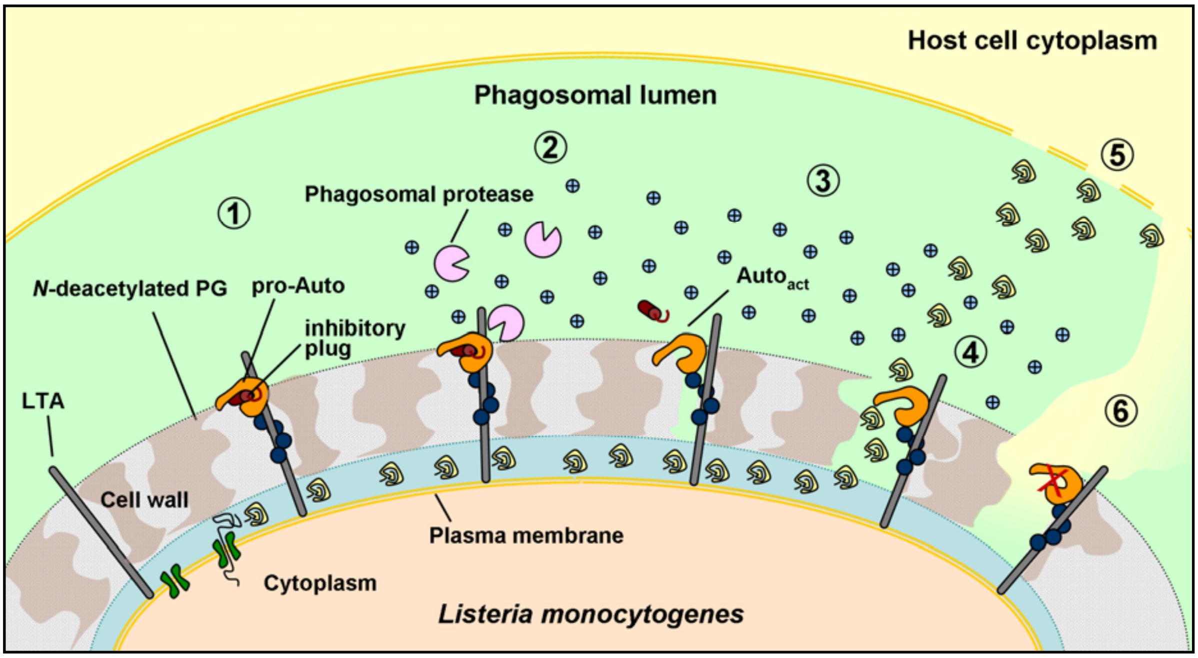


Fig. 5-3: Model of the role of Auto during phagosomal escape. (1) Inhibited pro-Auto is attached to lipoteichoic acid (LTA) in the cell wall. **(2)** A phagosomal protease cleaves pro-Auto within the *activation region*, releasing the inhibitory plug (red) and generating Auto$_{act}$. **(3)** Auto$_{act}$ has high activity in the acidified phagosomal environment (blue dots indicate H^+). **(4)** The autolytic action of Auto$_{act}$ accelerates secretion of previously retained factors. *N*-deacetylated cell wall regions and GW-mediated immobilization restrict zones of hydrolysis. **(5)** Secreted virulence factors promote efficient membrane lysis and phagosomal escape. **(6)** pH-neutralization upon phagosomal membrane lysis inactivates Auto.

5.1.6 Biotechnological potential of Auto

Cell wall degrading enzymes are valuable tools in medicine, food industry and agriculture, and are used in biotechnology to isolate intracellular bacterial components like nucleic acids or proteins (Ezaki *et al.*, 1990; Niwa *et al.*, 2005; Salazar & Asenjo, 2007).

The specificity of phage endolysins allow them to be used as targeted antibacterial agents to combat individual bacterial species (Borysowski *et al.*, 2006; Fischetti, 2005). The broad substrate range of Auto, by contrast, could be advantageous in applications requiring general bactericidal activity, such as surface decontamination, preservation of food and animal feed, as well as to lyse bacterial cells to isolate intracellular constituents. Auto could be particularly suitable due to the ease of large-scale production of soluble protein in *E. coli* (in contrast to HEWL, Schlörb *et al.*, 2005). Its protease resistance and heat stability would, moreover, allow its incorporation in large-scale formulations such as powders or capsules, mostly produced under high pressures and temperatures. An acidic pH-optimum compared to the optimum at neutral pH of HEWL could, further, make it a useful alternative in cases where HEWL is not sufficiently efficient.

The biotechnical use of Auto thus appears feasible. Expanded tests on the bactericidal activity – especially against lysozyme-resistant species such as pathogenic *Staphylococci* as well as Gram-negative bacteria – will have to follow in order to determine the applicability of Auto as a commercial antimicrobial agent.

5.2 Internalin-family LRR proteins

5.2.1 Cysteine and asparagine ladders

Asparagine or asparagine-cysteine ladders, consisting of regularly positioned asparagines or alternating asparagines and cysteines within an otherwise hydrophobic protein core, are a hallmark of LRR- and parallel β-helix-proteins (Kobe & Deisenhofer, 1995) and serve to stabilize proteins by a continuous array of enthalpically stable hydrogen bonds (Tsai *et al.*, 2005). Such ladders are, correspondingly, important for early stages of protein folding (Walter, 2002).

As an extracellular protein, InlJ is exposed to an oxidizing aerobic environment, especially during secretion. Perhaps surprisingly, its constituent cysteines remain reduced and are not oxidized to disulphides. The assumption, that an oxidizing environment would result in misfolding of cysteine-containing LRR proteins (Kajava, 1998), therefore, does not hold for InlJ. In the porcine RI, by contrast, an aerobic environment rapidly induces oxidation of the sulphhydryl groups, destroying its tertiary structure and resulting in its degradation (Blazquez *et al.*, 1996; Fominaya & Hofsteenge, 1992). The regular arrangement of hydrophobic cysteines in InlJ allows for their optimal integration into the hydrophobic core through van der Waals interactions (3.7 Å), while the polarizability of sulphur results in repeated S··H-C hydrogen bonds that additionally stabilize the protein enthalpically.

A further interesting feature of InlJ involves a second short asparagine ladder in LRRs 12 to 15. Note that both the cysteine ladder and the ‘classical’ asparagine ladder do not extend into these repeats. The asparagines creating this second asparagine ladder are located at repeat position 18, in the pre-β-strand turn and are presumably required to stabilize the C-terminal section of the InlJ LRR-domain (Fig. 5-4A).

Searching known structures of LRR and parallel β-helix proteins for comparable cysteine ladders, identifies four reduced cysteines in the parallel β-helix tailspike protein of phage SF6 (Barbirz, 2006) as the only related system. In another repeat protein, the β-helix antifreeze protein TmAFP from *Tenebrio molitor*, the situation is structurally unrelated to the Cys-ladder of InlJ in that positionally conserved cysteines within the 12-residue repeats create a row of intra-repeat disulphide bonds that substitute for a hydrophobic core (Liou *et al*., 2000, Fig. 5-4B). Although cysteines are commonly found within the core of most proteins, it appears that a highly regular system, as provided by the LRRs in InlJ, is required to benefit from the stabilizing effect of a cysteine ladder without incurring the risk of oxidation-induced protein unfolding.

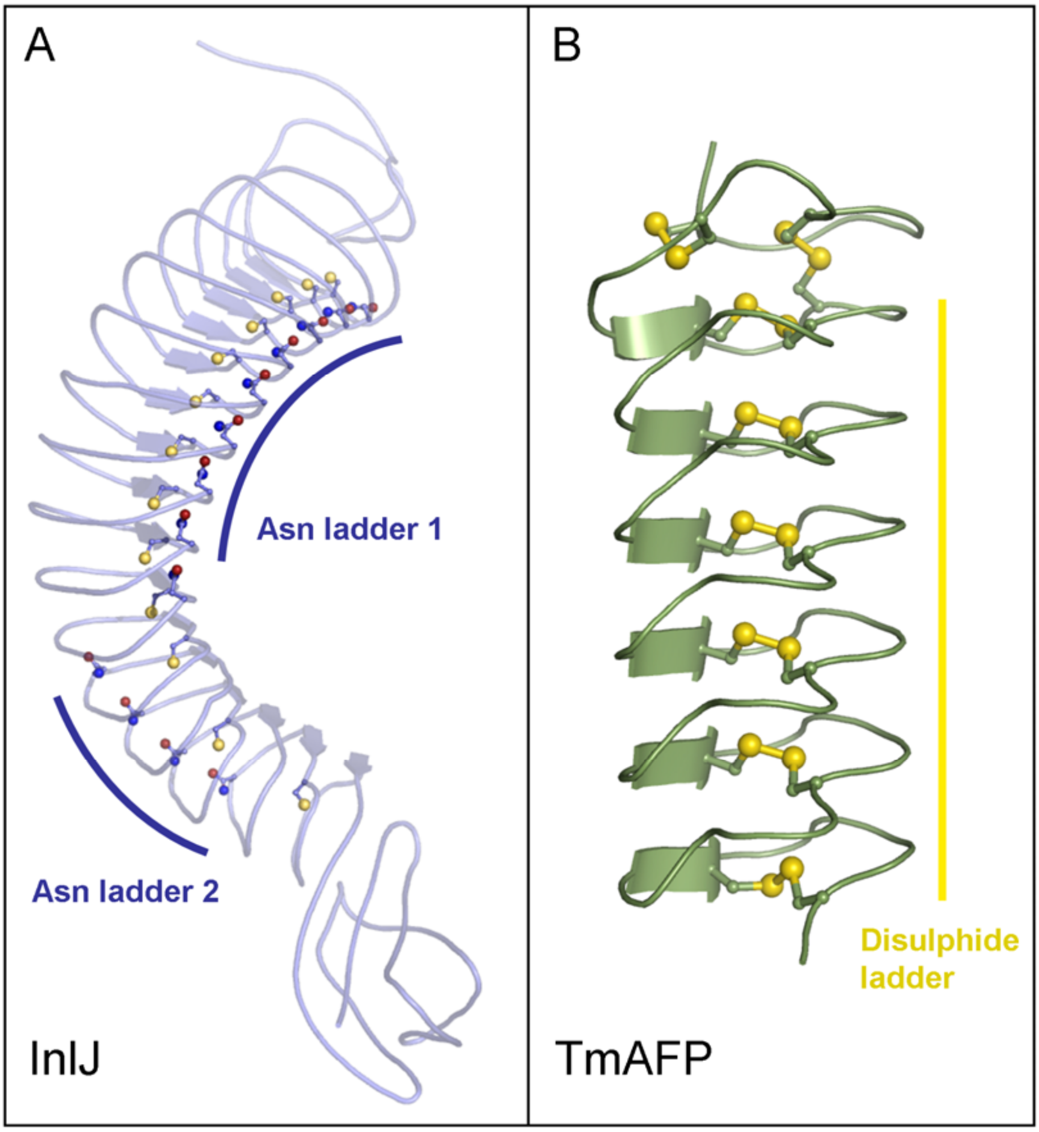


Fig. 5-4: Cysteine and asparagine ladders in InlJ' and disulphide ladder in AFP. (A) The 'classical' internalin asparagine ladder in repeat position 8 runs parallel to the cysteine ladder up to LRR10. From LRR11 on, where the cysteine ladder ceases, a second, shorter asparagine ladder in position 18 starts and continues up to LRR14. **(B)** In the antifreeze protein TmAFP from *Tenebrio molitor*, a disulphide ladder substitutes for a hydrophobic core.

5.2.2 LRR protein geometry

The overall curvature and twist of previously solved LRR-protein structures have generally been compared qualitatively. Some standardized geometric analyses have been undertaken especially to quantify the twist of the β-sheets as well as the radius and angle of protein curvature (Enkhbayar *et al*., 2004; McEwan *et al*., 2006). However, a consistent description of the LRR-geometry, taking into account the interdependence of the three rotation angles - curvature, twist and lateral bending - required to describe the overall geometry, has not been published.

Curvature

The average curving angles of the analysed proteins broadly confirm the hypothesis that curvature of LRR-proteins depends on the length of the LRR and hence on the secondary structure adopted by the variable region on the convex side of the LRR opposite the β-strand (Kobe & Deisenhofer, 1994). Amphipathic α-helices on the convex side of RI cause the repeats to adopt a more pronounced wedge-shape, while 3_{10}-helices in most internalins (except InlJ) and non-helical stretches in InlJ and YopM result in smaller curving angles.

However, the secondary structure of the variable region alone does not dictate the curving angle. This is illustrated by the fact that the curving angles for InlJ and YopM are roughly equivalent to those of all other internalin family members, despite the 3_{10}-helices of the latter being physically broader than the extended loops of InlJ and YopM. Clearly, additional effects such as the embedded ‘strings’ of water molecules (see above) may compensate for the lack of thickness of extended loops compared to 3_{10}-helices. It thus appears that curving angles of less than ~10° per repeat are incompatible with LRR-proteins.

Twist

β-Sheets, whether parallel, antiparallel or mixed, do not conform to the mathematical definition of parallel lines. Instead β-sheets, buried within proteins, have long been known to inherently twist in a clockwise direction when viewed along the β-sheet plane (Chothia, 1973). Examining the twist values of LRR-proteins in more detail reveals that the twist is not strictly unidirectional. It is thus not primarily dictated by the length of the repeat but rather by the constituent residues and their ability to integrate amongst the residues of neighbouring repeats. Interestingly, the twist of InlA in complex with E-cadherin does not differ substantially from uncomplexed InlA, indicating that ligand binding does not significantly affect the geometry of InlA.

Factors that affect the twist of LRR-proteins are not always clear. However, single residue insertions and/or deletions are the main cause for dramatic deviations from average twist values. In InlA, a twist angle above average correlates with the deletion of one residue in the affected LRR (Fig. 4-39B) and in InlJ, twist angles below average correlate with the insertion of an additional residue. This influence on the twist appears to be restricted to changes in the loop following the β-sheet, because insertions in other parts of the LRR, such as in the convex region of LRR4, 6 and 8 of YopM or LRR1 and 12 in InlJ do not obviously affect the twist angles. Clearly, disruptions of the regular LRR-pattern locally disturbs the packing efficiency giving rise to local deviations from normal twists.

Lateral bending

The degree to which individual β-strands within the β-sheet of LRR-proteins run parallel to each other within the "idealized" β-sheet plane has not been reported. Deviations from ideal parallelity would result in a lateral bending, which is normally difficult to discern by eye due to the more pronounced twist of the β-sheet.

Interestingly, the cumulative curves for lateral bending of the proteins analysed are similar in shape to that for the cumulative twist, indicating a degree of correlation. A reason for this could be the gain of "vertical" flexibility due to the "horizontal" twist between two β-strands.

5.2.3 Structural properties and evolution of LRR-proteins

A large number of leucine-rich-repeat proteins have been shown to engage their partners involving their inner, concave surface of the curved LRR-domain. Well-studied examples include InlA/hEC1 (Schubert *et al.*, 2002), InlB/Met (Machner *et al.*, 2003; Niemann *et al.*, 2007) and RI/RNAseA (Kobe & Deisenhofer, 1996). A host receptor for InlJ has presently not been identified, making co-crystallization experiments or activity assays to identify the regions of InlJ involved in protein-protein interactions currently impossible. The significant attenuation of a Δ*inlJ L. monocytogenes* mutant strain *in vivo*, however, implies a function in virulence (Sabet *et al.*, 2005) and indicates that the protein must in some way interact with a binding partner. The structure of InlJ reveals numerous solvent-exposed aromatic and hydrophobic side chains at the concave surface, presuming this region to be involved in binding a specific receptor (Fig. 4-31). Modelling *P. gingivalis* InlJ on InlJ of *L. monocytogenes*, places seven aromatic and/or hydrophobic residues within the concave surface region implying an analogous mode of receptor binding. Similar surface properties apply to InlG. Despite a backbone structure closely related to other family members, InlG presents a unique pattern of putative binding interface residues to the solvent at its concave side. The diversity of exposed surfaces among the internalins is clearly visible in their electrostatic surface potentials (Fig. 5-5).

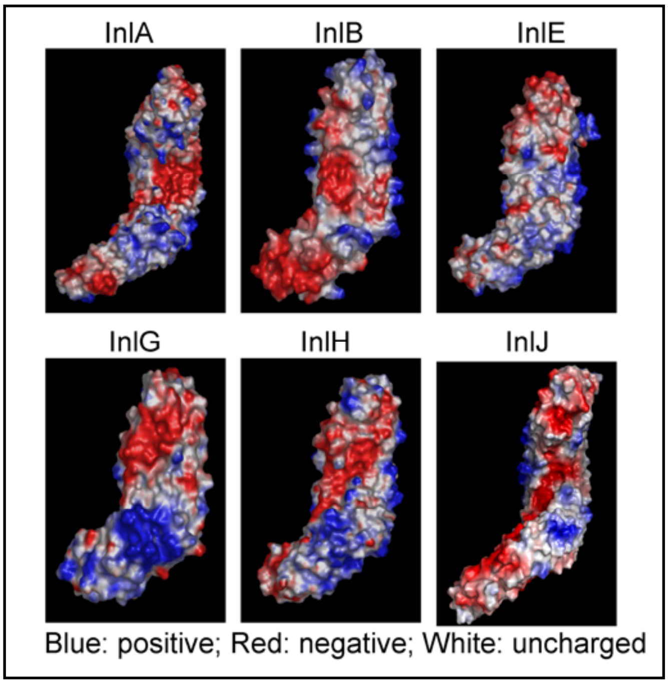


Fig. 5-5: Electrostatic surface potentials of different internalin family members, on the concave side of the LRR region. Although the architecture is conserved, each protein has a distinct distribution of its electrostatic surface potential, in particular with respect to the putative ligand-binding site of the LRR concave surface.

The distinct surface properties of each internalin family member indicate that these proteins are not the by-product of simple gene duplications. Instead, each appears to have evolved to achieve its own unique function. The best-characterized function of LRR proteins is the specific binding of receptor proteins. Internalins of *L. monocytogenes* have, accordingly, been proposed to have distinct evolutionary histories (Tsai *et al*., 2006). Horizontal transfer of internalin gene fragments between different *L. monocytogenes* lineages as well as intrachromosomal recombination presumably contributed to the evolution of the internalin genes (Tsai *et al*., 2006; Vazquez-Boland *et al*., 2001a). The repetitive juxtaposition of super-motifs encompassing two to three LRRs in RI and several small proteoglycans such as decorin indicates that LRR-elongation may result from duplication of such ancestral motifs (Matsushima *et al*., 2000; McEwan *et al*., 2006). In listerial internalins no such super-motifs have, however, been identified. The number of repeats in a specific internalin is, furthermore, invariant in different listerial strains, underlining the selective pressure exerted by their distinct functions. The consistent repeat sequence and length of LRRs in InlJ, distinct from that of other family members, suggests that InlJ may have evolved by repeating a shortened ancestral LRR rather than through gene exchange with other internalins. Fascinatingly, however, the first and last LRR of InlJ (adjacent to the neighbouring cap and Ig-subdomains) conform in length to those of other internalin family proteins. InlJ may thus have evolved from a protein of three or four LRRs, of which the first and last derive from a classical internalin. Once the central 21-residue repeat(s) had evolved with the characteristic cysteine, this repeat would have been multiplied to achieve the complete 15-LRR protein.

6 Conclusions and Outlook

6.1 The autolysin Auto

The link between bacterial autolysins and pathogenesis is only gradually being established. This study provides a biochemical and structural characterization of the listerial enzyme Auto, revealing its broad substrate range and its classification as an *N*-acetylglucosaminidase, two features probably shared by all members of the large GH73 family. Auto catalysis was described revealing an acidic pH-optimum and moderate substrate inhibition with one or several non-productive modes of substrate binding.

The crystal structure of the inactive pro-enzyme state of Auto was solved, revealing its mechanism of autoinhibition. The α-helix α0, an N-terminal extension of the otherwise lysozyme-like structure, was found to constitute a physical plug that blocks the substrate binding groove. It shields the 'catalytic' glutamate from bulk solvent until the enzyme is activated by cleaving an extending protease-sensitive loop, releasing the plug. A more detailed mutational study of the binding mode of the autoinhibitory plug could establish the contributions of single residues to overall binding affinity, revealing why the plug should be released upon cleavage of the distant flexible activation region.

Comparing Auto_{act} to other lysozyme-related enzymes allows a likely enzymatic mechanism to be proposed. The absence of a second carboxyl near the 'catalytic' Glu122 in the active site, allows the enzymatic mechanisms for Auto_{act} to be narrowed down to three alternatives:

i) a distant glutamate such as the conserved Glu156 could act as a base activating a water molecule that would nucleophilically attack the anomeric C1 of NAG from the α-side. This scenario would result in the inversion of configuration and is dependent upon substantial conformational rearrangements following substrate binding,

ii) substrate-assisted catalysis involving the carbonyl oxygen of the C2-*N*-acetamido group, also termed 'anchimeric assistance', with a covalent oxazoline intermediate, favouring retention of configuration of C1,

iii) substrate-assisted catalysis involving a nucleophilic carboxylate from the peptide bridge of the PG substrate. Depending on the distance of this substrate-derived nucleophile to the NAG-C1 atom, this would lead to either α- or β-anomeric cleavage products.

Experimentally determining the α- or β-conformation of Auto_{act} products would thus provide valuable information to rule out at least one these three hypotheses.

Very recent experiments (by Lilia Polle as part of her diploma thesis) indicate that Glu156 constitutes the second carboxylate group, with the result that the mutant E156Q is catalytically inactive. This result supports mechanism i) and implies that the protein undergoes a substantial conformational change upon ligand binding. The crystal structure of Auto_{act} in complex with a substrate analogue or inhibitor will allow the mode of substrate binding and hydrolysis to be described more fully.

Fluorescence-based tracking experiments *in vitro*, in combination with the structure-based elucidation of its physiological mode of activation and catalysis should allow the elucidation of the involvement of Auto in listerial phagosomal escape.

6.2 Internalin-family LRR proteins

The internalin domain of InlJ, solved in this study, represents the first report of a cysteine-rich LRR protein structure that includes a ladder of reduced cysteine residues, stacked in a regular array within the hydrophobic core of the molecule. This feature is not unique to InlJ, but occurs in several other prokaryotic proteins as well as in two entomopoxvirus proteins.

Experimental studies to assess the contribution of the cysteine ladder to protein stability and folding kinetics, though potentially informative, are difficult because the exchange of cysteine by any other amino acid will inevitably affect the geometry of the structure in multiple ways. In the geometric analysis, both Cys- and Asn-ladders, as well as amino acid insertions and deletions are observed to affect the curvature, twist and lateral bending of LRR proteins.

Structurally, there are clear indications for a receptor binding function of InlJ, consistent with results from a study of Δ*inlJ* mutants of *L. monocytogenes*. Several exposed hydrophobic and aromatic side chains on the concave surface of InlJ support the interpretation that this is an interface for recognition of proteins, as demonstrated for several other LRR proteins. The identification of the host receptor of InlJ is the next logical step in analysing and understanding the explicit function of InlJ in the infection process of *L. monocytogenes*.

Comparison of the surface properties of all known internalin structures indicate that, although all family members may have evolved from duplications of a common ancestor repeat, there has been a divergent development of altered surface properties, suitable for binding novel receptor proteins. Individual binding interfaces through unique surface residue combinations on the concave side of each internalin, but based on a common structural scaffold, means the internalin family is akin to a bacterial ‘multi-tool’ – with an easily adjustable binding surface to induce downstream signalling as required for infection.

The potential intracellular function of internalin proteins was also addressed in this work by heterologously expressing GFP-tagged internalins in human cells. However, a specific subcellular localization or a morphological modification of the transfected cells was not observed, supporting the idea that internalins function to specifically recognize extracellular receptors.

This study provides suitable expression vectors to produce recombinant GFP-InlX fusion proteins to be used in *in vitro* assays with potential target cell lines, as well as for interaction screens such as pull-down assays, in order to identify the receptors for the remaining internalins. This should allow the distinctive functions of the internalin ‘multi-tool’ in listerial pathophysiology to be described more comprehensively.

References

Adams, P. D., Gopal, K., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R. *et al.* (2004). Recent developments in the PHENIX software for automated crystallographic structure determination. *J.Synchrotron.Radiat.* **11**: 53-55.

Altekruse, S. F., Cohen, M. L., and Swerdlow, D. L. (1997). Emerging foodborne diseases. *Emerg.Infect.Dis.* **3**: 285-293.

Amano, K., Hayashi, H., Araki, Y., and Ito, E. (1977). The action of lysozyme on peptidoglycan with *N*-unsubstituted glucosamine residues. Isolation of glycan fragments and their susceptibility to lysozyme. *Eur.J.Biochem.* **76**: 299-307.

Andersen, M. D., Jensen, A., Robertus, J. D., Leah, R., and Skriver, K. (1997). Heterologous expression and characterization of wild-type and mutant forms of a 26 kDa endochitinase from barley (*Hordeum vulgare L.*). *Biochem.J.* **322**: 815-822.

Archibald, A. R., Hancock, I. C., and Harwood, C. R. (1993) Cell wall structure, synthesis, and turnover. In *Bacillus subtilis and other gram-positive bacteria: biochemistry, physiology, and molecular genetics.* Sonenshein, A. L., Hoch, J. A., Losick, R. (eds) pp 381-410. Washington, D.C.: American Society for Microbiology

Arnheim, N., Inouye, M., Law, L., and Laudin, A. (1973). Chemical studies on the enzymatic specificity of goose egg white lysozyme. *J.Biol.Chem.* **248**: 233-236.

Barbirz, S. (2006). Konservierte Struktur bei genetischer Mosaizität: die Tailspike Proteine dreier Phagen der Familie Podoviridae. *PhD Thesis*

Bergmann, B., Raffelsbauer, D., Kuhn, M., Götz, M., Hom, S. *et al.* (2002). InlA- but not InlB-mediated internalization of *Listeria monocytogenes* by non-phagocytic mammalian cells needs the support of other internalins. *Mol.Microbiol.* **43**: 557-570.

Birren, B., Lander, E., Galagan, J., Nusbaum, C., Devon, K. *et al.* (2006). The genome sequence of *Listeria monocytogenes* strain 10403S. *Direct submission, Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA*

Blazquez, M., Fominaya, J. M., and Hofsteenge, J. (1996). Oxidation of sulfhydryl groups of ribonuclease inhibitor in epithelial cells is sufficient for its intracellular degradation. *J.Biol.Chem.* **271**: 18638-18642.

Boneca, I. G., Dussurget, O., Cabanes, D., Nahori, M. A., Sousa, S. *et al.* (2007). A critical role for peptidoglycan *N*-deacetylation in *Listeria* evasion from the host innate immune system. *Proc.Natl.Acad.Sci.U.S.A* **104**: 997-1002.

Borysowski, J., Weber-Dabrowska, B., and Gorski, A. (2006). Bacteriophage endolysins as a novel class of antibacterial agents. *Exp.Biol.Med.(Maywood.)* **231**: 366-377.

Braun, L., Dramsi, S., Dehoux, P., Bierne, H., Lindahl, G. *et al.* (1997). InlB: an invasion protein of *Listeria monocytogenes* with a novel type of surface association. *Mol.Microbiol.* **25**: 285-294.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P. *et al.* (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr.D.Biol.Crystallogr.* **54**: 905-921.

Buchanan, S. G. and Gay, N. J. (1996). Structural and functional diversity in the leucine-rich repeat family of proteins. *Prog.Biophys.Mol.Biol.* **65**: 1-44.

Buist, G., Kok, J., Leenhouts, K. J., Dabrowska, M., Venema, G. *et al.* (1995). Molecular cloning and nucleotide sequence of the gene encoding the major peptidoglycan hydrolase of *Lactococcus lactis*, a muramidase needed for cell separation. *J.Bacteriol.* **177**: 1554-1563.

Cabanes, D., Dehoux, P., Dussurget, O., Frangeul, L., and Cossart, P. (2002). Surface proteins and the pathogenic potential of *Listeria monocytogenes*. *Trends Microbiol.* **10**: 238-245.

Cabanes, D., Dussurget, O., Dehoux, P., and Cossart, P. (2004). Auto, a surface associated autolysin of *Listeria monocytogenes* required for entry into eukaryotic cells and virulence. *Mol.Microbiol.* **51**: 1601-1614.

Camilli, A., Tilney, L. G., and Portnoy, D. A. (1993). Dual roles of plcA in *Listeria monocytogenes* pathogenesis. *Mol.Microbiol.* **8**: 143-157.

Capestany, C. A., Kuboniwa, M., Jung, I. Y., Park, Y., Tribble, G. D. *et al.* (2006). Role of the *Porphyromonas gingivalis* InlJ protein in homotypic and heterotypic biofilm development. *Infect.Immun.* **74**: 3002-3005.

Carroll, S. A., Hain, T., Technow, U., Darji, A., Pashalidis, P. *et al.* (2003). Identification and characterization of a peptidoglycan hydrolase, MurA, of *Listeria monocytogenes*, a muramidase needed for cell separation. *J.Bacteriol.* **185**: 6801-6808.

Chipman, D. M. and Sharon, N. (1969). Mechanism of lysozyme action. *Science* **165**: 454-465.

Chothia, C. (1973). Conformation of twisted beta-pleated sheets in proteins. *J.Mol.Biol.* **75**: 295-302.

Cibik, R. and Chapot-Chartier, M. P. (2004). Characterisation of autolytic enzymes in *Lactobacillus pentosus*. *Lett.Appl.Microbiol.* **38**: 459-463.

Clark, N. A., Lunacek, J. H., and Benedek, G. B. (1970). A Study of Brownian Motion Using Light Scattering. *Amer.J.Phys.* **38**: 575-585.

Coutinho, P. M. and Henrissat, B. (1999) Carbohydrate-active enzymes: an integrated database approach. In *Recent Advances in Carbohydrate Bioengineering*, Gilbert, H. J., Davies, G., Henrissat, B., Svensson, B. (eds) pp 3-12. Cambridge: The Royal Society of Chemistry

Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J. *et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**: W375-W383.

de Chastellier, C. and Berche, P. (1994). Fate of *Listeria monocytogenes* in murine macrophages: evidence for simultaneous killing and survival of intracellular bacteria. *Infect.Immun.* **62**: 543-553.

Derewenda, Z. S. (2004). Rational protein crystallization by mutational surface engineering. *Structure.* **12**: 529-535.

Dijkstra, A. J. and Keck, W. (1996). Peptidoglycan as a barrier to transenvelope transport. *J.Bacteriol.* **178**: 5555-5562.

Djinovic-Carugo, K., Gautel, M., Ylanne, J., and Young, P. (2002). The spectrin repeat: a structural platform for cytoskeletal protein assemblies. *FEBS Lett.* **513**: 119-123.

Domann, E., Wehland, J., Rohde, M., Pistor, S., Hartl, M. *et al.* (1992). A novel bacterial virulence gene in *Listeria monocytogenes* required for host cell microfilament interaction with homology to the proline-rich region of vinculin. *EMBO J.* **11**: 1981-1990.

Doyle, R. J., Chaloupka, J., and Vinter, V. (1988). Turnover of cell walls in microorganisms. *Microbiol.Rev.* **52**: 554-567.

Dramsi, S., Biswas, I., Maguin, E., Braun, L., Mastroeni, P. *et al.* (1995). Entry of *Listeria monocytogenes* into hepatocytes requires expression of InlB, a surface protein of the internalin multigene family. *Mol.Microbiol.* **16**: 251-261.

Eckert, C., Lecerf, M., Dubost, L., Arthur, M., and Mesnage, S. (2006). Functional analysis of AtlA, the major *N*-acetylglucosaminidase of *Enterococcus faecalis*. *J.Bacteriol.* **188**: 8513-8519.

Eckert, C., Magnet, S., and Mesnage, S. (2007). The *Enterococcus hirae* Mur-2 enzyme displays *N*-acetylglucosaminidase activity. *FEBS Lett.* **581**: 693-696.

Edman, P. and Begg, G. (1967). A protein sequenator. *Eur.J.Biochem.* **1**: 80-91.

Ehinger, J. (2002). Reinigung und Versuche zur Kristallisation von Listeriolysin O und Internalin-ähnlichen Proteinen aus *Listeria monocytogenes*. *Diploma Thesis*

Ellman (1958). A colorimetric method for determining low concentrations of mercaptans. *Arch.Biochem.Biophys.* **74**: 443-450.

Emsley, P. and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr.D.Biol.Crystallogr.* **60**: 2126-2132.

Engelbrecht, F., Dominguez-Bernal, G., Hess, J., Dickneite, C., Greiffenberg, L. *et al.* (1998). A novel PrfA-regulated chromosomal locus, which is specific for *Listeria ivanovii*, encodes two small, secreted internalins and contributes to virulence in mice. *Mol.Microbiol.* **30**: 405-417.

Enkhbayar, P., Kamiya, M., Osaki, M., Matsumoto, T., and Matsushima, N. (2004). Structural principles of leucine-rich repeat (LRR) proteins. *Proteins* **54**: 394-403.

Ezaki, T., Saidi, S. M., Liu, S. L., Hashimoto, Y., Yamamoto, H. *et al*. (1990). Rapid procedure to determine the DNA base composition from small amounts of gram-positive bacteria. *FEMS Microbiol.Lett.* **55**: 127-130.

Fischetti, V. A. (2005). Bacteriophage lytic enzymes: novel anti-infectives. *Trends Microbiol.* **13**: 491-496.

Fleming, A. (1929). On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. *Br.J.Exp.Pathol.* **10**: 226-236.

Fominaya, J. M. and Hofsteenge, J. (1992). Inactivation of ribonuclease inhibitor by thiol-disulfide exchange. *J.Biol.Chem.* **267**: 24655-24660.

Foster, S. J. (1995). Molecular characterization and functional analysis of the major autolysin of *Staphylococcus aureus* 8325/4. *J.Bacteriol.* **177**: 5723-5725.

Fukamizo, T. (2000). Chitinolytic enzymes: catalysis, substrate binding, and their application. *Curr.Protein Pept.Sci.* **1**: 105-124.

Gaillard, J. L., Berche, P., Frehel, C., Gouin, E., and Cossart, P. (1991). Entry of *L. monocytogenes* into cells is mediated by internalin, a repeat protein reminiscent of surface antigens from gram-positive cocci. *Cell* **65**: 1127-1141.

Gaillard, J. L., Berche, P., Mounier, J., Richard, S., and Sansonetti, P. (1987). In vitro model of penetration and intracellular growth of *Listeria monocytogenes* in the human enterocyte-like cell line Caco-2. *Infect.Immun.* **55**: 2822-2829.

Gellin, B. G. and Broome, C. V. (1989). Listeriosis. *JAMA* **261**: 1313-1320.

Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A. *et al*. (2001). Comparative genomics of *Listeria* species. *Science* **294**: 849-852.

Goebl, M. and Yanagida, M. (1991). The TPR snap helix: a novel protein repeat motif from mitosis to transcription. *Trends Biochem.Sci.* **16**: 173-177.

Grütter, M. G., Weaver, L. H., and Matthews, B. W. (1983). Goose lysozyme structure: an evolutionary link between hen and bacteriophage lysozymes? *Nature* **303**: 828-831.

Guerrero, S. A., Hecht, H. J., Hofmann, B., Biebl, H., and Singh, M. (2001). Production of selenomethionine-labelled proteins using simplified culture conditions and generally applicable host/vector systems. *Appl.Microbiol.Biotechnol.* **56**: 718-723.

Hamon, M., Bierne, H., and Cossart, P. (2006). *Listeria monocytogenes*: a multifaceted model. *Nat.Rev.Microbiol.* **4**: 423-434.

Hess, J., Gentschev, I., Szalay, G., Ladel, C., Bubert, A. *et al*. (1995). *Listeria monocytogenes* p60 supports host cell invasion by and in vivo survival of attenuated *Salmonella typhimurium*. *Infect.Immun.* **63**: 2047-2053.

Hillig, R. C. and Renault, L. (2006). Detecting and overcoming hemihedral twinning during the MIR structure determination of Rna1p. *Acta Crystallogr.D.Biol.Crystallogr.* **62**: 750-765.

Hof, H. (2001). *Listeria monocytogenes*: a causative agent of gastroenteritis? *Eur.J.Clin.Microbiol.Infect.Dis.* **20**: 369-373.

Höltje, J. V. and Tomasz, A. (1975). Lipoteichoic acid: a specific inhibitor of autolysin activity in *Pneumococcus*. *Proc.Natl.Acad.Sci.U.S.A* **72**: 1690-1694.

Horsburgh, G. J., Atrih, A., Williamson, M. P., and Foster, S. J. (2003). LytG of *Bacillus subtilis* is a novel peptidoglycan hydrolase: the major active glucosaminidase. *Biochemistry* **42**: 257-264.

Huard, C., Miranda, G., Wessner, F., Bolotin, A., Hansen, J. *et al.* (2003). Characterization of AcmB, an *N*-acetylglucosaminidase autolysin from *Lactococcus lactis*. *Microbiology* **149**: 695-705.

Kajava, A. V. (1998). Structural diversity of leucine-rich repeat proteins. *J.Mol.Biol.* **277**: 519-527.

Kajimura, J., Fujiwara, T., Yamada, S., Suzawa, Y., Nishida, T. *et al.* (2005). Identification and molecular characterization of an *N*-acetylmuramyl-L-alanine amidase Sle1 involved in cell separation of *Staphylococcus aureus*. *Mol.Microbiol.* **58**: 1087-1101.

Kamisango, K., Saiki, I., Tanio, Y., Okumura, H., Araki, Y. *et al.* (1982). Structures and biological activities of peptidoglycans of *Listeria monocytogenes* and *Propionibacterium acnes*. *J.Biochem.(Tokyo)* **92**: 23-33.

Kawamura, S., Ohno, K., Ohkuma, M., Chijiiwa, Y., and Torikata, T. (2006). Experimental verification of the crucial roles of Glu73 in the catalytic activity and structural stability of goose type lysozyme. *J.Biochem.(Tokyo)* **140**: 75-85.

Kobe, B. and Deisenhofer, J. (1994). The leucine-rich repeat: a versatile binding motif. *Trends Biochem.Sci.* **19**: 415-421.

Kobe, B. and Deisenhofer, J. (1995). Proteins with leucine-rich repeats. *Curr.Opin.Struct.Biol.* **5**: 409-416.

Kobe, B. and Deisenhofer, J. (1996). Mechanism of ribonuclease inhibition by ribonuclease inhibitor protein based on the crystal structure of its complex with ribonuclease A. *J.Mol.Biol.* **264**: 1028-1043.

Koch, J. and Stark, K. (2006). Significant increase of listeriosis in Germany--epidemiological patterns 2001-2005. *Euro.Surveill* **11**: 85-88.

Kocks, C., Gouin, E., Tabouret, M., Berche, P., Ohayon, H. *et al.* (1992). *L. monocytogenes*-induced actin assembly requires the actA gene product, a surface protein. *Cell* **68**: 521-531.

Krissinel, E. and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr.D Biol.Crystallogr.* **60**: 2256-2268.

Kuhn, M. and Goebel, W. (1989). Identification of an extracellular protein of *Listeria monocytogenes* possibly involved in intracellular uptake by mammalian cells. *Infect.Immun.* **57**: 55-61.

Kuroki, R., Weaver, L. H., and Matthews, B. W. (1999). Structural basis of the conversion of T4 lysozyme into a transglycosidase by reengineering the active site. *Proc.Natl.Acad.Sci.U.S.A* **96**: 8949-8954.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J.Appl.Cryst.* **26**: 283-291.

Liou, Y. C., Tocilj, A., Davies, P. L., and Jia, Z. (2000). Mimicry of ice structure by surface hydroxyls and water of a beta-helix antifreeze protein. *Nature* **406**: 322-324.

Loessner, M. J., Kramer, K., Ebel, F., and Scherer, S. (2002). C-terminal domains of *Listeria monocytogenes* bacteriophage murein hydrolases determine specific recognition and high-affinity binding to bacterial cell wall carbohydrates. *Mol.Microbiol.* **44**: 335-349.

Loessner, M. J., Wendlinger, G., and Scherer, S. (1995). Heterogeneous endolysins in *Listeria monocytogenes* bacteriophages: a new class of enzymes and evidence for conserved holin genes within the siphoviral lysis cassettes. *Mol.Microbiol.* **16**: 1231-1241.

Low, L. Y., Yang, C., Perego, M., Osterman, A., and Liddington, R. C. (2005). Structure and lytic activity of a *Bacillus anthracis* prophage endolysin. *J.Biol.Chem.* **280**: 35433-35439.

Machner, M. P., Frese, S., Schubert, W. D., Orian-Rousseau, V., Gherardi, E. *et al*. (2003). Aromatic amino acids at the surface of InlB are essential for host cell invasion by *Listeria monocytogenes*. *Mol.Microbiol.* **48**: 1525-1536.

Malcolm, B. A., Rosenberg, S., Corey, M. J., Allen, J. S., de, B. A. *et al*. (1989). Site-directed mutagenesis of the catalytic residues Asp-52 and Glu-35 of chicken egg white lysozyme. *Proc.Natl.Acad.Sci.U.S.A* **86**: 133-137.

Marino, M., Banerjee, M., Jonquieres, R., Cossart, P., and Ghosh, P. (2002). GW domains of the *Listeria monocytogenes* invasion protein InlB are SH3-like and mediate binding to host ligands. *EMBO J.* **21**: 5623-5634.

Marino, M., Braun, L., Cossart, P., and Ghosh, P. (1999). Structure of the InlB leucine-rich repeats, a domain that triggers host cell invasion by the bacterial pathogen *L. monocytogenes*. *Mol.Cell* **4**: 1063-1072.

Marquis, H., Doshi, V., and Portnoy, D. A. (1995). The broad-range phospholipase C and a metalloprotease mediate listeriolysin O-independent escape of *Listeria monocytogenes* from a primary vacuole in human epithelial cells. *Infect.Immun.* **63**: 4531-4534.

Marquis, H., Goldfine, H., and Portnoy, D. A. (1997). Proteolytic pathways of activation and degradation of a bacterial phospholipase C during intracellular infection by *Listeria monocytogenes*. *J.Cell Biol.* **137**: 1381-1392.

Matsumura, I. and Kirsch, J. F. (1996). Is aspartate 52 essential for catalysis by chicken egg white lysozyme? The role of natural substrate-assisted hydrolysis. *Biochemistry* **35**: 1881-1889.

Matsushima, N., Ohyanagi, T., Tanaka, T., and Kretsinger, R. H. (2000). Super-motifs and evolution of tandem leucine-rich repeats within the small proteoglycans--biglycan, decorin, lumican, fibromodulin, PRELP, keratocan, osteoadherin, epiphycan, and osteoglycin. *Proteins* **38**: 210-225.

McCarter, J. D. and Withers, S. G. (1994). Mechanisms of enzymatic glycoside hydrolysis. *Curr.Opin.Struct.Biol.* **4**: 885-892.

McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005). Likelihood-enhanced fast translation functions. *Acta Crystallogr.D.Biol.Crystallogr.* **61**: 458-464.

McEwan, P. A., Scott, P. G., Bishop, P. N., and Bella, J. (2006). Structural correlations in the family of small leucine-rich repeat proteins and proteoglycans. *J.Struct.Biol.* **155**: 294-305.

McLauchlin, J., Mitchell, R. T., Smerdon, W. J., and Jewell, K. (2004). *Listeria monocytogenes* and listeriosis: a review of hazard characterisation for use in microbiological risk assessment of foods. *Int.J.Food Microbiol.* **92**: 15-33.

McLaughlan, A. M. and Foster, S. J. (1997). Characterisation of the peptidoglycan hydrolases of *Listeria monocytogenes* EGD. *FEMS Microbiol.Lett.* **152**: 149-154.

Mead, P. S., Slutsker, L., Dietz, V., McCaig, L. F., Bresee, J. S. *et al.* (1999). Food-related illness and death in the United States. *Emerg.Infect.Dis.* **5**: 607-625.

Mengaud, J., Ohayon, H., Gounon, P., Mege, R.-M., and Cossart, P. (1996). E-cadherin is the receptor for internalin, a surface protein required for entry of *L. monocytogenes* into epithelial cells. *Cell* **84**: 923-932.

Meyrand, M., Boughammoura, A., Courtin, P., Mezange, C., Guillot, A. *et al.* (2007). Peptidoglycan *N*-acetylglucosamine deacetylation decreases autolysis in *Lactococcus lactis*. *Microbiology* **153**: 3275-3285.

Milohanic, E., Jonquieres, R., Cossart, P., Berche, P., and Gaillard, J. L. (2001). The autolysin Ami contributes to the adhesion of *Listeria monocytogenes* to eukaryotic cells via its cell wall anchor. *Mol.Microbiol.* **39**: 1212-1224.

Mosavi, L. K., Cammett, T. J., Desrosiers, D. C., and Peng, Z. Y. (2004). The ankyrin repeat as molecular architecture for protein recognition. *Protein Sci.* **13**: 1435-1448.

Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr.D.Biol.Crystallogr.* **53**: 240-255.

Nelson, K. E., Fouts, D. E., Mongodin, E. F., Ravel, J., Deboy, R. T. *et al.* (2004). Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen *Listeria monocytogenes* reveal new insights into the core genome components of this species. *Nucleic Acids Res.* **32**: 2386-2395.

Neuvirth, H., Raz, R., and Schreiber, G. (2004). ProMate: a structure based prediction program to identify the location of protein-protein binding sites. *J.Mol.Biol.* **338**: 181-199.

Niemann, H. H., Jäger, V., Butler, P. J., van den, H. J., Schmidt, S. *et al.* (2007). Structure of the human receptor tyrosine kinase met in complex with the *Listeria* invasion protein InlB. *Cell* **130**: 235-246.

Niwa, T., Kawamura, Y., Katagiri, Y., and Ezaki, T. (2005). Lytic enzyme, labiase for a broad range of Gram-positive bacteria and its application to analyze functional DNA/RNA. *J.Microbiol.Methods* **61**: 251-260.

Ooi, A., Hussain, S., Seyedarabi, A., and Pickersgill, R. W. (2006). Structure of internalin C from *Listeria monocytogenes*. *Acta Crystallogr.D.Biol.Crystallogr.* **62**: 1287-1293.

Otwinowski, Z. and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol.* Carter, C. W. and Sweet, R. M.(ed) pp 307-326. Academic Press

Pamer, E. G. (2004). Immune responses to *Listeria monocytogenes*. *Nat.Rev.Immunol.* **4**: 812-823.

Pape, T. and Schneider, T. R. (2004). HKL2MAP: a graphical user interface for phasing with SHELX programs. *J.Appl.Cryst.* **37**: 843-844.

Pentecost, M., Otto, G., Theriot, J. A., and Amieva, M. R. (2006). *Listeria monocytogenes* invades the epithelial junctions at sites of cell extrusion. *PLoS.Pathog.* **2**: e3-

Pizarro-Cerda, J., Jonquieres, R., Gouin, E., Vandekerckhove, J., Garin, J. *et al*. (2002). Distinct protein patterns associated with *Listeria monocytogenes* InlA- or InlB-phagosomes. *Cell Microbiol.* **4**: 101-115.

Popowska, M. (2004). Analysis of the peptidoglycan hydrolases of *Listeria monocytogenes*: multiple enzymes with multiple functions. *Pol.J.Microbiol.* **53 Suppl**: 29-34.

Popowska, M. and Markiewicz, Z. (2004). Murein-hydrolyzing activity of flagellin FlaA of *Listeria monocytogenes*. *Pol.J.Microbiol.* **53**: 237-241.

Portnoy, D. A., Chakraborty, T., Goebel, W., and Cossart, P. (1992). Molecular determinants of *Listeria monocytogenes* pathogenesis. *Infect.Immun.* **60**: 1263-1267.

Raffelsbauer, D., Bubert, A., Engelbrecht, F., Scheinpflug, J., Simm, A. *et al*. (1998). The gene cluster inlC2DE of *Listeria monocytogenes* contains additional new internalin genes and is important for virulence in mice. *Mol.Gen.Genet.* **260**: 144-158.

Reichelt, J., Dieterich, G., Kvesic, M., Schomburg, D., and Heinz, D. W. (2005). BRAGI: linking and visualization of database information in a 3D viewer and modeling tool. *Bioinformatics.* **21**: 1291-1293.

Riddles, P. W., Blakeley, R. L., and Zerner, B. (1983). Reassessment of Ellman's reagent. *Methods Enzymol.* **91**: 49-60.

Sabet, C., Lecuit, M., Cabanes, D., Cossart, P., and Bierne, H. (2005). LPXTG protein InlJ, a newly identified internalin involved in *Listeria monocytogenes* virulence. *Infect.Immun.* **73**: 6912-6922.

Salazar, O. and Asenjo, J. A. (2007). Enzymatic lysis of microbial cells. *Biotechnol.Lett.* **29**: 985-994.

Schindler, M., Mirelman, D., and Sharon, N. (1977). Substrate-induced evolution of lysozymes. *Biochim.Biophys.Acta* **482**: 386-392.

Schleifer, K. H. and Kandler, O. (1972). Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol.Rev.* **36**: 407-477.

Schlörb, C., Ackermann, K., Richter, C., Wirmer, J., and Schwalbe, H. (2005). Heterologous expression of hen egg white lysozyme and resonance assignment of tryptophan side chains in its non-native states. *J.Biomol.NMR* **33**: 95-104.

Schneider, T. R. and Sheldrick, G. M. (2002). Substructure solution with SHELXD. *Acta Crystallogr.D.Biol.Crystallogr.* **58**: 1772-1779.

Schubert, W. D., Göbel, G., Diepholz, M., Darji, A., Kloer, D. *et al*. (2001). Internalins from the human pathogen *Listeria monocytogenes* combine three distinct folds into a contiguous internalin domain. *J.Mol.Biol.* **312**: 783-794.

Schubert, W. D. and Heinz, D. W. (2003). Structural aspects of adhesion to and invasion of host cells by the human pathogen *Listeria monocytogenes*. *Chembiochem.* **4**: 1285-1291.

Schubert, W. D., Urbanke, C., Ziehm, T., Beier, V., Machner, M. P. *et al*. (2002). Structure of internalin, a major invasion protein of *Listeria monocytogenes*, in complex with its human receptor E-cadherin. *Cell* **111**: 825-836.

Schwarzenbacher, R., Godzik, A., Grzechnik, S. K., and Jaroszewski, L. (2004). The importance of alignment accuracy for molecular replacement. *Acta Crystallogr.D.Biol.Crystallogr.* **60**: 1229-1236.

Seeliger, H. P. R. and Jones, D. (1986) Genus *Listeria*. In *Bergey's Manual of Systematic Bacteriology*, Holt, J. G. (ed) pp 1235-1245. Baltimore, MD: Williams and Wilkins

Sheldrick, G. M. (2002). Macromolecular phasing with SHELXE. *Z.Kristallogr.* **217**: 644-650.

Sheldrick, G. M. (2004). SHELXC. *University of Göttingen, Germany*

Sheldrick, G. M. and Schneider, T. R. (1997). SHELXL: High-resolution refinement. *Methods Enzymol.* **277**: 319-343.

Shen, Y., Naujokas, M., Park, M., and Ireton, K. (2000). InlB-dependent internalization of *Listeria* is mediated by the Met receptor tyrosine kinase. *Cell* **103**: 501-510.

Shockman, G. D. and Höltje, J. V. (1994) Microbial peptidoglycan (murein) hydrolases. In *Bacterial Cell Wall*, Ghuysen, J. M., Hakenbeck, R. (eds) pp 131-166. Amsterdam: Elsevier Science B.V.

Smith, G. A., Marquis, H., Jones, S., Johnston, N. C., Portnoy, D. A. *et al*. (1995). The two distinct phospholipases C of *Listeria monocytogenes* have overlapping roles in escape from a vacuole and cell-to-cell spread. *Infect.Immun.* **63**: 4231-4237.

Smith, T. J., Blackman, S. A., and Foster, S. J. (2000). Autolysins of *Bacillus subtilis*: multiple enzymes with multiple functions. *Microbiology* **146 (Pt 2)**: 249-262.

Snyder, A. and Marquis, H. (2003). Restricted translocation across the cell wall regulates secretion of the broad-range phospholipase C of *Listeria monocytogenes*. *J.Bacteriol.* **185**: 5953-5958.

Steen, A., Buist, G., Horsburgh, G. J., Venema, G., Kuipers, O. P. *et al*. (2005). AcmA of *Lactococcus lactis* is an *N*-acetylglucosaminidase with an optimal number of LysM domains for proper functioning. *FEBS J.* **272**: 2854-2868.

Stenesh, J. (1984) *Experimental Biochemistry.* Allyn and Bacon, Boston

Storoni, L. C., McCoy, A. J., and Read, R. J. (2004). Likelihood-enhanced fast rotation functions. *Acta Crystallogr.D.Biol.Crystallogr.* **60**: 432-438.

Taylor, H. O. and Leslie, A. G. W. (1998). A program to detwin merohedrally twinned data. *CCP4 Newsl.Prot.Crystallogr.* **35**: 9-

Terwisscha van Scheltinga, A. C., Armand, S., Kalk, K. H., Isogai, A., Henrissat, B. *et al*. (1995). Stereochemistry of chitin hydrolysis by a plant chitinase/lysozyme and X-ray structure of a complex with allosamidin: evidence for substrate assisted catalysis. *Biochemistry* **34**: 15619-15623.

Thibault, F., Langowski, J., and Leberman, R. (1992). Optimizing protein crystallization by aggregate size distribution analysis using dynamic light scattering. *J.Cryst.Growth* **122**: 50-59.

Thunnissen, A. M., Dijkstra, A. J., Kalk, K. H., Rozeboom, H. J., Engel, H. *et al*. (1994). Doughnut-shaped structure of a bacterial muramidase revealed by X-ray crystallography. *Nature* **367**: 750-753.

Thunnissen, A. M., Isaacs, N. W., and Dijkstra, B. W. (1995a). The catalytic domain of a bacterial lytic transglycosylase defines a novel class of lysozymes. *Proteins* **22**: 245-258.

Thunnissen, A. M., Rozeboom, H. J., Kalk, K. H., and Dijkstra, B. W. (1995b). Structure of the 70-kDa soluble lytic transglycosylase complexed with bulgecin A. Implications for the enzymatic mechanism. *Biochemistry* **34**: 12729-12737.

Ton-That, H., Marraffini, L. A., and Schneewind, O. (2004). Protein sorting to the cell wall envelope of Gram-positive bacteria. *Biochim.Biophys.Acta* **1694**: 269-278.

Tsai, H. H., Reches, M., Tsai, C. J., Gunasekaran, K., Gazit, E. *et al*. (2005). Energy landscape of amyloidogenic peptide oligomerization by parallel-tempering molecular dynamics simulation: significant role of Asn ladder. *Proc.Natl.Acad.Sci.U.S.A* **102**: 8174-8179.

Tsai, Y. H., Orsi, R. H., Nightingale, K. K., and Wiedmann, M. (2006). *Listeria monocytogenes* internalins are highly diverse and evolved by recombination and positive selection. *Infect.Genet.Evol.* **6**: 378-389.

Tweten, R. K. and Iandolo, J. J. (1983). Transport and processing of staphylococcal enterotoxin B. *J.Bacteriol.* **153**: 297-303.

Vazquez-Boland, J. A., Dominguez-Bernal, G., Gonzalez-Zorn, B., Kreft, J., and Goebel, W. (2001a). Pathogenicity islands and virulence evolution in *Listeria*. *Microbes.Infect.* **3**: 571-584.

Vazquez-Boland, J. A., Kuhn, M., Berche, P., Chakraborty, T., Dominguez-Bernal, G. *et al.* (2001b). *Listeria* pathogenesis and molecular virulence determinants. *Clin.Microbiol.Rev.* **14**: 584-640.

Walter, M. (2002). Die parallele β-Helix der Pektat-Lyase aus *Bacillus subtilis*: Stabilität, Faltungsmechanismus und Faltungsmutanten. *PhD Thesis*

Weaver, L. H., Grütter, M. G., and Matthews, B. W. (1995). The refined structures of goose lysozyme and its complex with a bound trisaccharide show that the "goose-type" lysozymes lack a catalytic aspartate residue. *J.Mol.Biol.* **245**: 54-68.

Whitcomb, D. C. and Lowe, M. E. (2007). Human pancreatic digestive enzymes. *Dig.Dis.Sci.* **52**: 1-17.

Wollert, T., Heinz, D. W., and Schubert, W. D. (2007a). Thermodynamically reengineering the listerial invasion complex InlA/E-cadherin. *Proc.Natl.Acad.Sci.U.S.A* **104**: 13960-13965.

Wollert, T., Pasche, B., Rochon, M., Deppenmeier, S., van den, H. J. *et al.* (2007b). Extending the host range of *Listeria monocytogenes* by rational protein design. *Cell* **129**: 891-902.

Yeates, T. O. (1997). Detecting and overcoming crystal twinning. *Methods Enzymol.* **276**: 344-358.

Zwart, P. H., Grosse-Kunstleve, R. W., and Adams, P. D. (2005). Xtriage and Fest: automatic assessment of x-ray data and substructure structure factor estimation. *CCP4 Newsletter Winter*

Danksagung

Endlich ein Platz um all denen zu danken, die mich während dieser Arbeit begleitet und zu ihrem Gelingen beigetragen haben!

Mein erster Dank geht an Dr. Wolf-Dieter Schubert, für ungezählte Stunden fachlicher Anleitung, inspirierender Diskussion und unerschöpflicher Hilfsbereitschaft bei allen Problemen und Fragen, die während der Arbeit auftraten. Die heimlich von mir durchgeführte Langzeitstudie, wie viel eine herkömmliche Biologin in einem Zeitraum von etwa 3.5 Jahren von einem einzigen Betreuer über Röntgenkristallographie lernen kann, ohne dass dieser jemals verzweifelt zusammenbricht, Amok läuft oder sich für immer im reziproken Raum versteckt, sprengt wirklich jede Statistik! Danke!

Prof. Dirk Heinz möchte ich herzlich für die Übernahme des Referats danken, vor allem aber dafür, dass er eines schönen Tages im Spätsommer 2003 neue Schuhe kaufen wollte... Sein anhaltendes Interesse, viele hilfreiche Tipps und besonders seine Begeisterungsfähigkeit für neue Ideen haben meine Arbeit sehr bereichert. Und nicht zuletzt danke ich ihm auch dafür, dass er immer genau dann am meisten an den Erfolg geglaubt hat, wenn ich am Zweifeln war. Vielen Dank dafür!

Herrn Prof. Michael Steinert danke ich sehr herzlich für die Übernahme des Korreferats.

Bei Prof. Ralf Mendel bedanke ich mich herzlich für die Übernahme des Prüfungsvorsitzes.

Christin Holland, Matze Haffke, Alexis Kuetchou und Lilia Polle schulde ich ein riesengroßes Dankeschön für ihren großen Arbeitseinsatz und ihre großartigen Beiträge zu Teilprojekten meiner Arbeit, beziehungsweise deren Weiterführung (nicht zu vergessen die unvergleichlichen Torten von Matze!).

Dr. Manfred Nimtz und seinem Team danke ich für seine fachkundige Unterstützung bei der Planung, Durchführung und Auswertung von massenspektrometrischen Experimenten, die dieser Arbeit und nicht zuletzt meinem methodischen Horizont sehr gut getan haben.

Dr. Theresia Stradal danke ich für ihre freundliche Unterstützung bei allen zellbiologischen Untersuchungen und für die Teilnahme an meinem Thesis-Komitee.

Auch an Dr. Lothar Jänsch geht ein großes Dankeschön, für viele bereichernde Hinweise im Thesis-Komitee.

Dr. Peter Schumann, Dr. Rüdiger Pukall und Jennifer Gregor von der DSMZ schulde ich großen Dank für die gute Zusammenarbeit und Ihre Hilfe bei der Arbeit mit Peptidoglykan.

Prof. Dr. George Sheldrick und Dr. Ina Dix von der Universität Göttingen möchte ich herzlich für Ihre unkomplizierte Hilfe in allen erdenklichen kristallographischen Nöten danken.

Bei Rita Getzlaff möchte ich mich herzlich für viele aufschlußreiche N-terminale Sequenzierungen bedanken, und Dr. Joop van den Heuvel möchte ich dafür danken, dass er mich so geduldig in die geheimnisvolle Welt der Klonierungen eingeführt hat.

Dr. Victor Wray und Dr. Hans-Jürgen Hecht verdienen auch ein großes Dankeschön, weil sie immer als Anlaufstelle für kleine und große Fragen – von Zwillings-Schizophrenie bis Anglizismen-Wahnsinn – für mich zur Stelle waren.

Die Durchführung dieser Arbeit wäre nicht möglich gewesen ohne die unzähligen kleinen und großen Hilfen meiner Kollegen aus der Arbeitsgruppe, die für ein fantastisches Arbeitsklima, lustige Mittagspausen und legendäre DoKo-Treffen gesorgt haben. Besonders zu erwähnen seien hier Katrin Bolte (Organisation legendärer Kohlfahrten, Kochabende, Fußball-Events…), Carina Büttner (unvergessliche Dienstreisen), Davide Ferraris (der “deutscheste” Italiener unter der Sonne – *grazie per i biscotti e espressi*), Hartmut Niemann (der Manuskript-Fuchs und konkurrenzlose Aufmunterer), Joachim Reichelt (der wahre Angulator), und natürlich das ‘Komische Zentrum’ bestehend aus Gregor Hagelüken (der Linux-Fuchs: “warte, ich hab da mal n’ Skript für geschrieben”…), Maike Rochon (die Lach-Muskelkater-nach-der-Mittagspause-Garantin), und (Infec)Thomas Wollert (der noch nie auf irgendwas keine Antwort wusste).

Sabine Schmidt und Ute Widow müsste ich eine ganze Seite extra widmen, nämlich dafür, dass sie mir als “Quereinsteigerin” wirklich ALLES von Anfang an gezeigt haben und nicht auch nur ein einziges Mal genervt auf dumme Fragen reagiert haben. Vielen, vielen Dank dafür, Euch zwei!

Ein extra Dankeschön geht natürlich auch noch an Jörg: die mittäglichen Musik-Fachsimpeleien, Songtext-Zitat-Sessions und mentalen Aufpäppeleien haben mir im letzten Jahr wirklich gefehlt!

An meine (z. T. Ex-) Bürogenossen Christian, Boris, Maria und Ulrich geht ein dicker Dank für ein ausgesprochen nettes „zweites Zuhause“, das mir sehr fehlen wird. Christian sahnt natürlich noch ein extra Dankeschön ab fürs penible Korrekturlesen, für ausgiebige J&W – Schwelgereien und seinen unvergleichlichen Humor. *What-ho…erm…something else!*

Ein riesengroßes Dankeschön schulde ich auch meinen Freunden Steph, Jessi, Diana & Harald, dafür, dass sie es nie aufgegeben haben, mich hin und wieder doch mal loszueisen und am „wirklichen Leben“ teilhaben zu lassen. Ohne das W:O:A, ausgiebige Samstags-Frühstücks-Zeremonien und angebranntes Kräuterbaguette vom Grill wäre ich vermutlich irgendwann ganz im reziproken Raum verschwunden.

Der bedingungslose Rückhalt meiner großartigen Familie hat den ganzen langen Weg hierher überhaupt erst möglich gemacht, und ich hoffe, Ihr wisst, wie viel es mir bedeutet, dass ihr immer an mich geglaubt habt, und dass trotzdem irgendwie immer noch alles genauso ist, wie früher. *¡Muchissimas gracias!*
Dem Carsten muss ich nochmal extra für seine Leidensfähigkeit und seine Argusaugen danken (wie um alles in der Welt erkennt man, dass ein Punkt tiefgestellt ist?!).

Und Sam, hier all die kleinen und großen Dinge aufzuzählen für die ich Dir unendlich dankbar bin, würde ein noch dickeres Buch als dieses füllen. Es sei an dieser Stelle nur eins gesagt: Machen wir uns nix vor - es gibt nur einen Melzerator auf der ganzen Welt ... DANKE!

Lebenslauf

Persönliche Daten

Maike Bublitz

geb. 05.05.1978 in Wolfenbüttel

ledig

deutsch

Schulische Ausbildung

1984 – 1986	Deutsche Fernschule, Gallito Ciego, Peru
1986 – 1988	Grundschule am Geitelplatz, Wolfenbüttel
1988 – 1990	Orientierungsstufe Wallstraße, Wolfenbüttel
1990 – 1997	Theodor-Heuss-Gymnasium, Wolfenbüttel Abschluss: Abitur

Akademische Ausbildung

1997 – 2003	Studium der Biologie an der Technischen Universität Carolo-Wilhelmina zu Braunschweig Abschluss: Diplom (Note: Sehr gut)
2002 – 2003	Diplomarbeit an der Bundesforschungsanstalt für Landwirtschaft, Braunschweig, Institut für Technologie und Biosystemtechnik: Verkapselung und Trocknung von *Pseudomonas fluorescens* zur biologischen Schädlingsbekämpfung
2003 – 2007	Promotion an der Gesellschaft für Biotechnologische Forschung (seit 2006: Helmholtz-Zentrum für Infektionsforschung), AG Molekulare Wirt-Pathogen Interaktionen im Bereich Strukturbiologie, Braunschweig